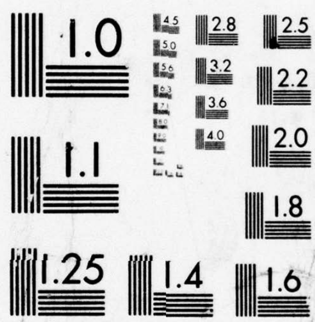
AD-A033 350 SYSTEMS CONTROL INC PALO ALTO CALIF F/G 1/2
SCIDNT I, THEORY AND APPLICATION.(U)
DEC 74 N K GUPTA, W E HALL N00014-72-C-0328
UNCLASSIFIED TR-3 NL

1 OF 1
AD
A033350

END
DATE
FILMED
2-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SYSTEMS CONTROL, INC.
1801 PAGE MILL ROAD
PALO ALTO, CALIFORNIA 94304
TELEX: 348433
TELEPHONE (415)
494-1165

(1)

(11) December 1974

Engineering Report (12) 76p.

ADA033350

(6) SCIDNT I,
THEORY AND APPLICATION

(9) Technical Report No. 3

(14) TR-3

Prepared for

Office of Naval Research
800 North Quincy Road
Arlington, Virginia 22217

and

Naval Air Test Center
Patuxent River, Maryland

Contract No. (15) N00014-72-C-0328

NR 215-213
code 210

Prepared by:
(10) N.K. Gupta
W.E. Hall

389333

1B

ABSTRACT

This technical report describes an application of the maximum likelihood approach for the identification of aircraft parameters. The theory and the computational aspects of the likelihood approach are discussed in detail.

A computer program is written to implement this techique. The details of the computer program are given in a separate document.

FORWORD

SCIDNT is a computer program for the identification of aircraft stability and control derivatives. This program was developed under financial sponsorship of the Naval Air Test Center. Mr. Roger Burton of NATC was technical monitor for this contract, N00014-72-C-0328.

Acknowledgement must be further made to the Office of Naval Research, under whose sponsorship many of the algorithms for this program were determined. Mr. David Siegel of ONR was technical monitor for this previous ONR work.

TABLE OF CONTENTS

I. INTRODUCTION

To meet current aircraft and rotorcraft flight testing requirements, there is a genuine need for an accurate procedure to evaluate stability and control coefficients from recorded response data. Most techniques developed in the past have either not had the capability to treat highly noisy data or could not be practically implemented within a computer program which is able to identify several aircraft parameters at a time. The work reported here and the resulting computer program, SCIDNT, is a development to fulfill this requirement.

Past methods have been mostly based on least square techniques (e.g., curve fitting or the equation error) or by approximating aircraft motions by low order systems to determine the frequency and damping of specific aircraft modes. The resultant algorithms required flight test time that may be excessive and only a few parameters can be identified. The basis of the present approach, on the other hand, is the likelihood principle of statistical inference theory. The likelihood function is defined and the parameter values are chosen to maximize this function. This approach leads to an algorithm which resolves many of the problems encountered by less comprehensive methods.

This report is a companion document to the User's Guide for SCIDNT [4], and provides the SCIDNT user with the theoretical foundations of the program. In particular, the optimization methods used, the Kalman filter used for identification in the presence of process noise, and the complete sensitivity equations are discussed. Examples of SCIDNT performance are also given.

The report is organized as follows. Chapter II discusses the maximum likelihood method and the numerical procedure which could be used to identify the aircraft stability and control coefficients from flight data. Chapter III presents the identification results from simulation data as well as flight data. Here, some of the data preprocessing algorithms which were used to reduce the flight data are also given. Chapter IV summarizes the report. An expression for the likelihood function is derived in Appendix A and Appendix B gives the aircraft equations of motion.

II. MAXIMUM LIKELIHOOD ESTIMATION OF AIRCRAFT PARAMETERS

2.1 INTRODUCTION

The maximum likelihood method is based on a fundamentally different statistical concept from least square methods for the identification of parameters from input-output data. Suppose it is possible to make a set of observations on a system, whose model has p unknown parameters θ. For any given set of values of the parameters θ from the feasible set Θ, we can assign a probability $p(z|\theta)$ to each outcome z. If the outcome of an actual experiment is z, it is of importance to know which sets of values of θ might have led to these observations. This concept is embedded in the likelihood function $\mathscr{L}(\theta|z)$. This function is of fundamental importance in estimation theory because of the likelihood principle of Fisher and others which states that if the system model is correct, all information about unknown parameters is contained in the likelihood function. The maximum likelihood method finds a set of parameters θ to maximize this likelihood function

$$\hat{\theta} = \max_{\theta \varepsilon \Theta} \mathscr{L}(\theta/z) \tag{2.1}$$

In other words, the probability of the outcome z is higher with parameters in the model than with any other values of parameters from the feasible set. Conceptually, this technique can be summarized as follows:

> "Find the probability density functions of the observations for all possible combinations of unknown parameter values. Select the density function whose value is highest among all density functions at the measured values of the observations. The corresponding parameter values are the maximum likelihood estimates."

Suppose θ can take three possible values: θ_1, θ_2, and θ_3. Let the probability density functions of observations z for these three values of θ be shown in Figure 2.1. Then, if the actual observation is z, θ_2 is the maximum likelihood estimate of θ.

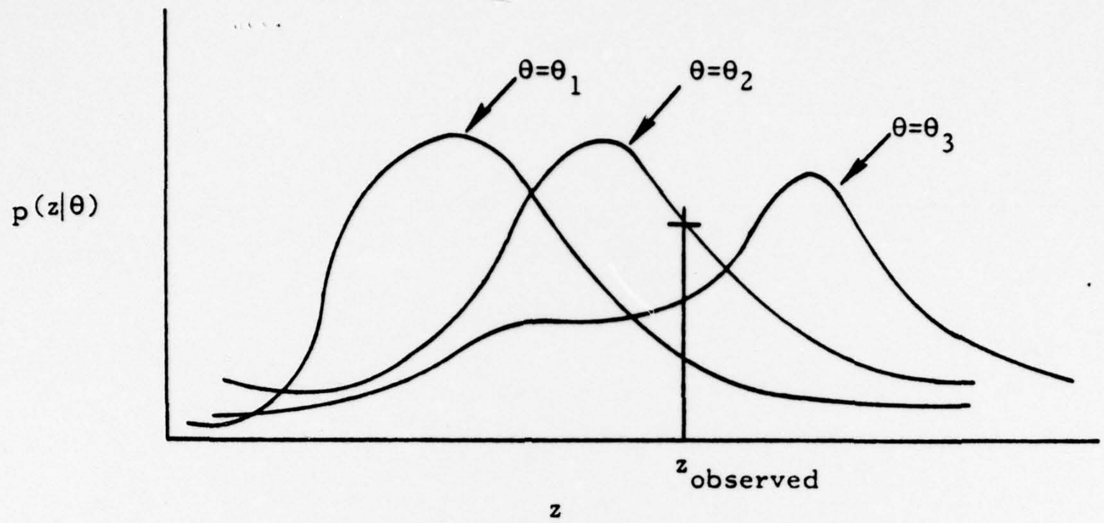


Figure 2.1 Maximum Likelihood Estimates

Usually, it is more convenient to work with the negative of the logarithm of the likelihood function (it is possible to do so because the logarithm is a monotonic function). Starting from a priori values, the parameters are updated so that the value of negative log-likelihood function J at the observed value of the outputs decreases continuously.

The great asset of the maximum likelihood method is that it can be used with linear or nonlinear models in the presence of process and measurement noise. The maximum likelihood estimates are asymptotically unbiased, consistent and efficient.

The details of the maximum likelihood identification procedure are given in Appendix A. A simplified flow chart is shown in Figure 2.2. Since the Kalman filter is a maximum likelihood approach to system state estimation, it is used to determine the state time history with fixed parameters. The negative log-likelihood function which depends on the innovations (difference between the observations and estimated value of the observations) and the covariance of the innovations can then be determined. The first and second

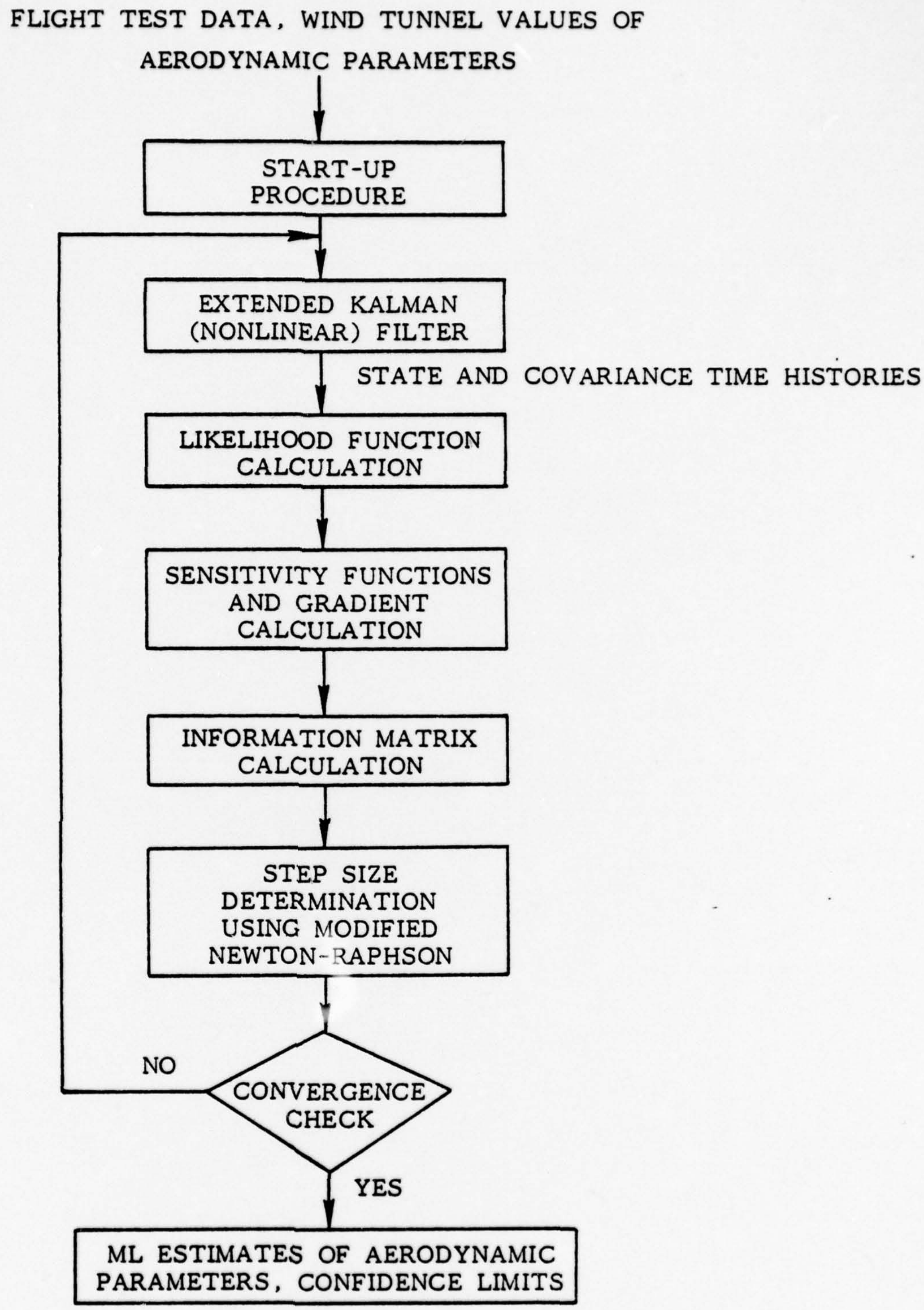


Figure 2.2 Flow Chart of Maximum Likelihood Identification Program

gradients of J require the computation of gradients of innovations and its covariance with respect to various unknown parameters. These, in turn, are determined by propagating state sensitivity equations. The parameter step size which reduces J is computed using modified Newton-Raphson or a Gauss Newton procedure. Notice that it is possible to consider noise statistics as parameters and identify them. A byproduct of this numerical procedure is the information matrix. The information matrix provides measures of the identifiability of the parameters. If this information matrix is singular for a certain excitation of the aircraft, then not all parameters in the model are identifiable from this maneuver. The eigenvalues and eigenvectors of this symmetric matrix give useful information about how well the parameters can be identified from given data.

2.2 MAXIMUM LIKELIHOOD IDENTIFICATION

There are two implementations of the likelihood method which have been used here.

(a) Output error, in which the noise in inputs and gust effects are small and can be neglected.

(b) Process noise, in which the input noise and/or gust effects are significant and must be included. The gust and input noise RMS values can be identified, if required.

In the output error implementation, the state equations are propagated directly to determine the estimated state vector at any time. The Kalman filter is not required because there are no unknown disturbances in the state equations which would cause error in the estimated value of the state. Therefore, the innovations and gradients of the innovations are determined directly by integrating the state equations and sensitivity equations. In the process noise implementation, the state equations are forced by unknown random inputs. A Kalman filter must, then, be used to determine the estimated state variable vector. The gradients of the innovations for system parameters

are computed by integrating the sensitivities of the estimated state in the Kalman filter equations. The process noise covariance can also be identified because the Kalman filter gains and innovations covariance depend on this parameter.

The maximum likelihood method is very general and many features can be added to any implementation. The following features are considered important for applications.

(i) Detailed instrument models are required. The measurements are assumed to have bias and random noise and correction terms for coupling with other states. In addition, the angle-of-attack and sideslip measurements usually have scale factor errors.

For example, the angle-of-attack measurement is

$$\alpha_m = k_\alpha \alpha + \frac{k\alpha \ell\alpha}{V} q + b\alpha + n\alpha \qquad (2.1)$$

where $k\alpha$ is the scale factor, $b\alpha$ is the bias and $n\alpha$ is random noise.

(ii) In most implementations, the measurement noise covariance matrix is estimated by

$$\hat{R} = \frac{1}{N} \sum_{i=1}^{N} \nu(i)\ \nu^T(i) \qquad (2.2)$$

where $\nu(i)$ are innovations. In general, for finite data length, this would give an estimate of measurement noise covariance matrix with all nonzero elements and no special structure. In other words, Eq. (2.2) assumes that there are $m(m+1)/2$ unknown elements in measurement noise statistics. If the structure of R is known, as is usually the case in aircraft applications, the number of unknown elements in R is fewer than $m(m+1)/2$. For example, if the measurement noise in different channels is independent, only diagonal terms of R need to be estimated. Using Eq. (2.2) will lead to an overparameterized model and incorrect parameter estimates.

The statistically correct estimate of measurement noise covariance matrix for independent measurement noise sources is,

$$\hat{R} = \text{diag}\ \frac{1}{N}\left\{\sum_{i=1}^{N} \nu(i)\ \nu^T(i)\right\} \tag{2.3}$$

Information Aggregation

Any available a priori information about parameters can be incorporated into the identification program. This is done by finding the weighted mean of the a priori parameter values and the output of the identification program, the weights varying inversely as the corrresponding covariances. This has been done before [1]. A similar approach can be used to combine estimates from different identification runs and thus make full use of available data.

In the vector case, let θ_1 and θ_2 be two different estimates of the parameter vector with information matrices M_1 and M_2 respectively. Then we can aggregate these two estimates to get a new information matrix and parameter vector.

$$M = M_1 + M_2 \tag{2.5}$$

$$\hat{\theta} = M^{-1}\ \{M_1\hat{\theta}_1 + M_2\hat{\theta}_2\} \tag{2.6}$$

2.3 COMPUTATIONAL ASPECTS OF MAXIMUM LIKELIHOOD ESTIMATES [2]

The application of maximum likelihood identification in practice requires efficient computational algorithms to maximize the likelihood function. It is not uncommon to find situations where the likelihood surface has multiple maxima, saddle-points, discontinuities and singular Hessian in the parameter space. The application of the steepest descent method leads to extremely slow convergence rate and the straightforward application of the Newton-Raphson and the Gauss-Newton methods may lead either to no convergence or convergence to wrong stationary points. From a statistical veiwpoint, only the maximum of the likelihood function provides an unbiased, consistent and efficient estimate. Thus, it is important to locate the true maximum of the

likelihood function. With the present computation methods of nonlinear programming, in general, this could be an extremely difficult and time-consuming task. In aircraft applications, a priori information about the parameters is used to choose a good starting value and impose physically meaningful constraints on the parameters. This increases the chances of approaching the absolute maxima.

2.3.1 Identifiability Problems and Solutions

Anomalies also occur in the likelihood function due to inadequate model specification and parameterization. Some of these anomalies that lead to numerical difficulties are:

1. Discontinuities and Singularities: There are examples where the likelihood function increases up to a maximum value and then falls immediately to minus infinity. In other cases, the likelihood function may rise suddenly to infinity for some particular parameter values. These problems can often be avoided either by changing the model or by changing the parameterization.

2. Singular or Nearly Singular Information Matrix: This is one of the most common problems in parameter identification since it is difficult to determine a priori whether a given sample has adequate information for estimating all the parameters in the model. Generally, the tendency in practical applications is to include all the parameters about which there is some uncertainty. This leads to overparameterization which in turn produces a singular or nearly singular information matrix. The likelihood surface in turn has long, curved and narrow ridges along which the convergence rate is extremely slow. Under-parameterization does not really provide the solution since it may lead to spurious local maxima and saddle-points.

The basic iteration in gradient-type nonlinear programming methods is

$$\theta_{i+1} = \theta_i - \rho_i R_i^{-1} g_i \tag{2.7}$$

where θ_i is the parameter vector at the i^{th} iteration, g_i is a vector of gradients of the negative log-likelihood function $J(\theta)$, i.e., the j^{th} component of g_i is given as

$$g_i(j) = \sum_{t=1}^{N} \left\{ \nu^T \beta^{-1} \frac{\partial \nu}{\partial \theta_i(j)} - \tfrac{1}{2} \nu^T \beta^{-1} \frac{\partial \beta}{\partial \theta_i(j)} \beta^{-1} \nu + \tfrac{1}{2} \operatorname{Tr} \left(\beta^{-1} \frac{\partial \beta}{\partial \theta_i(j)} \right) \right\} \tag{2.8}$$

R_i is an approximation to the second partial matrix

$$\left. \frac{\partial^2 J}{\partial \theta^2} \right|_{\theta = \theta_i}$$

where ρ_i is a scalar step size parameter chosen to ensure that $J(\theta_{i+1}) < J(\theta_i) - \varepsilon$ where ε is a positive number that can be chosen in a variety of ways. In the Gauss-Newton method used here, the matrix of second partials is approximated by

$$R_i(j,k) = M_i(j,k) \cong \sum_{t=1}^{N} \left\{ \left(\frac{\partial \nu^T}{\partial \theta_i(j)} \beta^{-1} \frac{\partial \nu}{\partial \theta_i(k)} \right) + \tfrac{1}{2} \operatorname{Tr} \left[\beta^{-1} \frac{\partial \beta}{\partial \theta_i(j)} \beta^{-1} \frac{\partial \beta}{\partial \theta_i(k)} \right] \right\} . \tag{2.9}$$

where M is the approximation of R. The first partials of innovations ν and their covariance can be computed by solving linear difference equations, i.e., the sensitivity equations. Notice that in this approximation computation of second partials of state and covariance is not required.

There are several problems with this method. As long as M_i^{-1} is positive semi-definite, it is possible to find a step which will reduce the cost. As mentioned before, most of the problems arise when the information matrix is almost singular, i.e., the eigenvalues are spread far apart. This near-singularity of the information matrix manifests itself in several ways.

(a) During the computation of the inverse of the information matrix, the positive-definiteness may be lost because of round-off errors.

(b) The step-size may be very large in those directions which correspond to small eigenvalues of the information matrix. Let

$$M = \sum_{i=1}^{m} \lambda_i v_i v_i^T \tag{2.10}$$

$$M^{-1} = \sum_{i=1}^{m} \frac{1}{\lambda_i} v_i v_i^T \tag{2.11}$$

$$\Delta\theta = -\alpha \sum_{i=1}^{m} \frac{1}{\lambda_i} v_i v_i^T g$$

$$= -\sum_{i=1}^{m} \alpha \left(\frac{v_i^T g}{\lambda_i}\right) v_i \tag{2.12}$$

where λ_i are the eigenvalues and v_i the corresponding eigenvectors of M. Thus, if $v_i^T g$ is not small for small eigenvalues λ_i, the step in the direction of v_i is large. Usually, for small eigenvalues, $v_i^T g$ is also small, but it is the difference of large and almost equal numbers. Therefore, in most cases, it is only a numerical problem. Complications arise from local irregularities in the likelihood function. Several remedies can be used, depending on the problem complexity.

(a) $\Delta\theta = -\alpha M^{-1} g$

Let $\Delta\theta' = M^{-1} g$ (2.13)

then $\Delta\theta = -\alpha\Delta\theta'$ (2.14)

Positive-definiteness of M could be lost during the inversion. This can be avoided by finding $\Delta\theta'$ by solving the following linear equations

$$M\Delta\theta' = g \tag{2.15}$$

The accuracy of Eq. (2.15) also depends on conditioning of M (ratio of maximum and minimum eigenvalues) but will be much superior as compared to Eq. (2.13).

(b) Conditioning of M can be improved. Premultiply both sides of Eq. (2.15) by A^{-1}, where

$$A = \text{diag}\left[\sqrt{M_{11}},\ \sqrt{M_{22}}\ .\ .\ .\ \sqrt{M_{mm}}\right] \tag{2.16}$$

to get

$$(A^{-1}MA^{-1})\ (A\Delta\theta') = A^{-1}g$$

or

$$M^*\Delta\theta^* = g^* \tag{2.17}$$

All elements of M^* are between -1 and +1. In general, this matrix is much better conditioned than M. Equation (2.17) can be used to solve for $\Delta\theta^*$ and then $\Delta\theta'$ is obtained by rescaling.

(c) The <u>rank deficient inverse</u> of M can also be used. In this procedure, the eigenvalues of matrix M are arranged in decreasing order of magnitude and all eigenvalues below a threshold b are neglected in using Eq. (2.11) for the inverse, i.e.,

$$\hat{M}^{-1} = \sum_{i=1}^{k} \frac{1}{\lambda_i} v_i v_i^T \tag{2.18}$$

where

$$\lambda_i \geq b \qquad\qquad i \leq k$$

$$\lambda_i < b \qquad\qquad i > k \tag{2.19}$$

This is equivalent to searching for the maximum of the likelihood function in the subspace spanned by $v_i (i \leq k)$. If this process is continued until the end, the estimates may never converge to their maximum likelihood values. It should also be noted that since only large eigenvalues are included, step $\Delta\theta$ may be "small".

(d) Instead of neglecting eigenvalues of M, which are smaller than a certain threshold b, these smaller eigenvalues are increased to b. Thus, the inverse of M is approximated as

$$\hat{M}^{-1} = \sum_{i=1}^{k} \frac{1}{\lambda_i} v_i v_i^T + \frac{1}{b} \sum_{i=k+1}^{m} v_i v_i^T \tag{2.20}$$

In this case, when the program converges, the likelihood function will reach a maxima. Convergence may be slow.

(e) This fix of the Gauss-Newton method is a two step procedure. The step is taken by finding M^{-1} using Eq. (2.18). Then, another step direction is computed using only the small eigenvalues, i.e.,

$$\Delta\theta_2 = -\alpha \sum_{i=k+1}^{m} \frac{v_i v_i^T g}{\lambda_i} \tag{2.21}$$

Then step size α is chosen by a one dimensional search. As the algorithm approaches the maximum, it may be worthwhile to carry out one dimensional search in each $v_i (i > k)$ direction.

The last approach when used with normalized information matrix M* of Eq. (2.17) is one of the best. It is recommended if the number of unknown parameters is large. In most cases, fix (d) is adequate.

III. EXAMPLES

3.1 INTRODUCTION

The computer program (SCIDNT) is based on the maximum likelihood method outlined in Chapter II and Appendix A. Many examples have been performed to test the usefulness and accuracy of the technique. These examples include both longitudinal and lateral motions and simulation and flight test data from a high performance swept-wing Navy fighter.

The equations of motion used in the identification are given in Appendix B. In addition to α, u, q and θ there is an additional state in the longitudinal case corresponding to the gust angle-of-attack, α_g. Similarly, there is a fifth state in the lateral equations of motion which corresponds to random fluctuations in sideslip angle because of lateral gusts.

3.2 SIMULATION DATA

Simulated data provides a tool for program checkout and experimentation for special problems (e.g., input design). In this section, the important simulated data runs by SCIDNT are summarized. A DC-8 simulation at 600 fps forward speed and sea level altitude was used both for longitudinal and lateral motions. The computer program PLANE generates the simulation data.

3.2.1 Longitudinal Motion

The values of parameters, used in the simulation to generate the data, are shown in Table 3.1 An elevator input is used to excite the aircraft [see Figure 3.1]. Five noisy instruments measure the resulting aircraft response. It is clear that only the short period mode is excited.

In the first case, it is assumed that there are no gust effects. SCIDNT is started using least square parameter values. Five parameters, $Z\alpha$, $M\alpha$, Mq, Z_{δ_e} and M_{δ_e}, which determine the short period motions are identified in addition to bias and random noise in instruments. The identified parameters are shown in Table 3.1. Also shown are the root mean square values of the

Table 3.1

DC-8 SIMULATED DATA

Velocity 600 fps
Altitude Sea Level

Parameter	True Value	Estimated Parameter Value (with 1σ bound)	
		Output Error	Process Noise
Z_α	-0.3237	-0.3168	-0.3205 (.0025)
Z_u	-1.074	*	*
Z_q	1.0	0.9212	1.0313 (.057)
γ_o	0.0	*	*
X_α	0.02845	*	*
X_u	-0.006386	*	*
X_q	0.0	*	*
M_α	-1.204	*	*
M_u	0.0	*	*
M_q	-0.3735	-0.5006	-0.402 (.012)
Z_{δ_e}	-0.0178	-0.131	0.0 (.056)
X_{δ_e}	0.0	*	*
M_{δ_e}	-1.381	-1.411	-1.396 (.014)
ω_c	0.5	*	*
Q	0.002	*	0.0024 (.00018)
k_α	1.0	*	*
$\frac{k_\alpha \ell_\alpha}{V}$	0.0	*	*
σ_α	0.005	*	0.011
σ_u	0.5	*	0.93
σ_q	0.005	*	0.0105
σ_θ	0.005	*	0.0102
σ_{a_z}	0.5	*	0.5

* Not identified

[All in units of ft., rad., sec.]

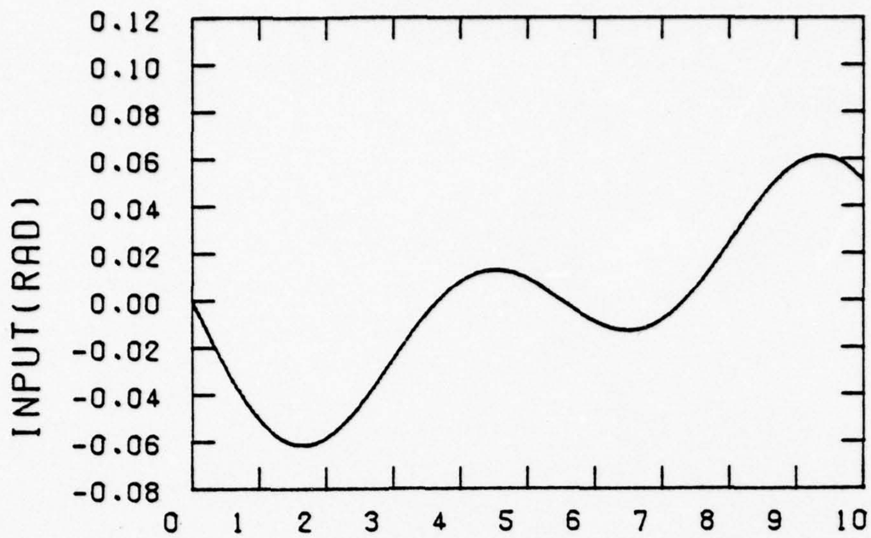


Figure 3.1 Elevator Input for DC-8 Simulation

random noise in instruments. The time history of the simulated and predicted aircraft response are shown in Figure 3.2. The fit is good. Next, a simulation data is generated with gust effects included. The pwoer spectral density of the white noise driving the gust state is .002. Five short period parameters and the power spectral density of the white noise are estimated from the simulation data. The time history fits are given in Figure 3.3 and the estimated values in Table 3.1. Also shown are one standard deviation values on parameter estimates.

It is important to understand the significance of the standard deviation on parameter estimation errors. Because of instrument errors and unknown gust effects, the estimated parameter values are not the same as true parameter values. In other words, there is an estimation error. Consider the parameter estimates to be random samples from a normally distributed population whose mean (unknown) is the true parameter value. Then the parameter estimates are within two standard deviations of the true parameter value about 95% times. Usually, the parameter estimates do not have a normal distribution and a relation such as the Chebyshev inequality must be used to

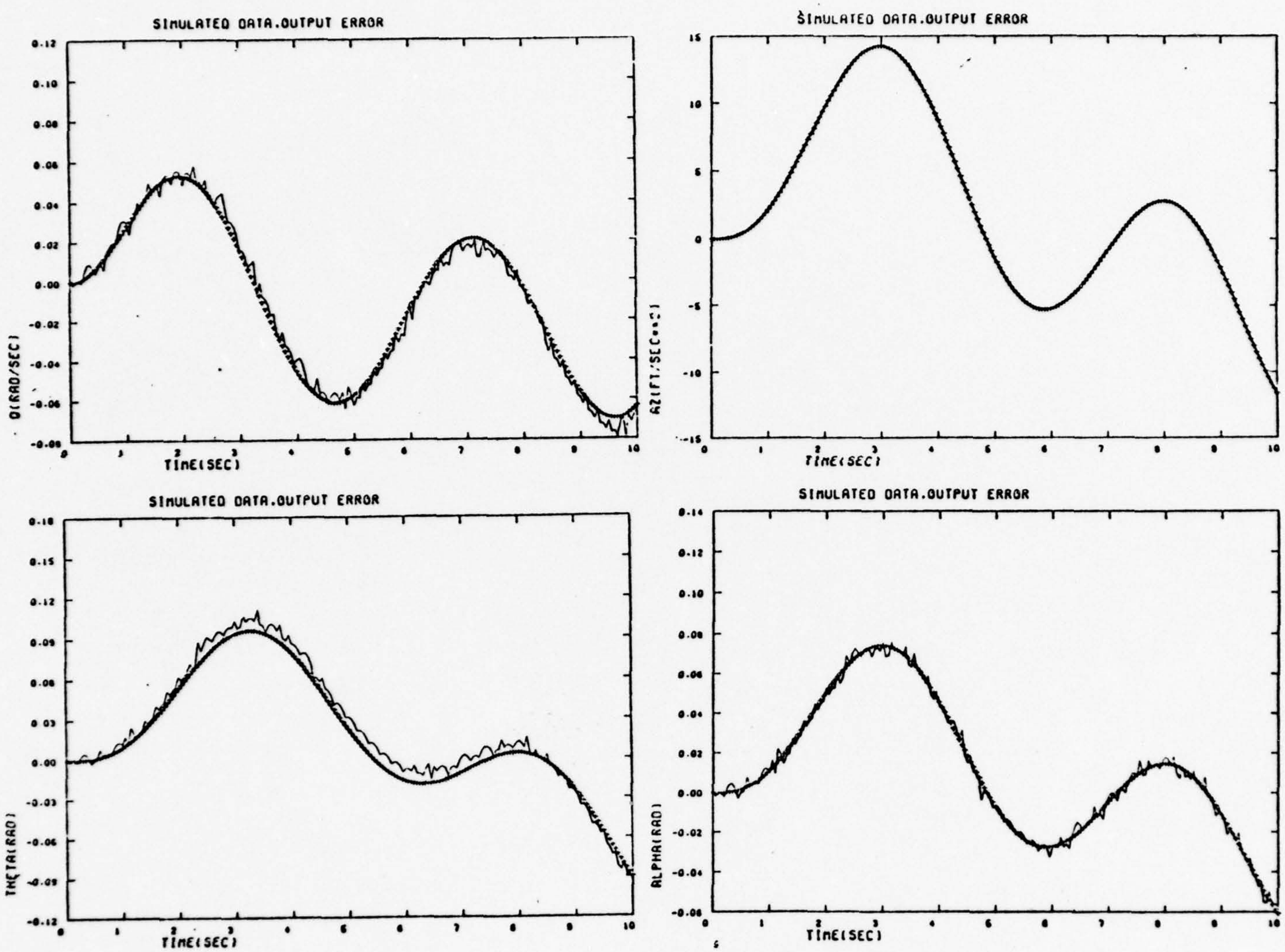


Figure 3.2 True and Predicted Measurements from Simulated Data Without Process Noise (DC-8)

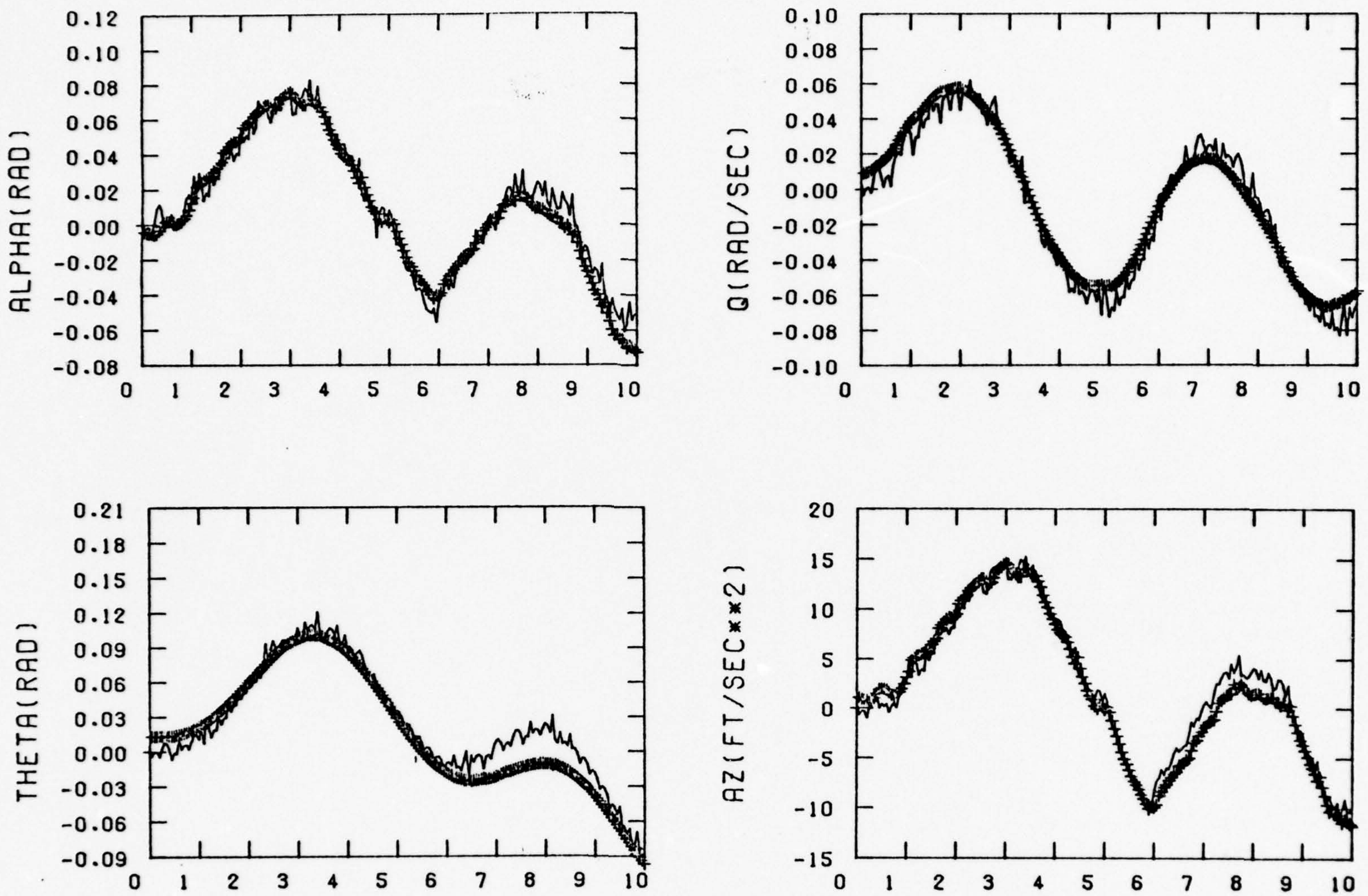


Figure 3.3 True and Predicted Measurements from Simulated Data With Process Noise (DC-8)

give more useful meaning to the standard deviation*. Notice that this is a probabilistic concept and does not ensure anything definite in a specific case. The maximum likelihood method, when used with the numerical procedure of Section 2, gives the standard deviations of parameter estimation errors as a byproduct of the parameter identification.

To explain the meaning of the estimation error standard deviation in this example, consider the parameter $Z\alpha$. Assume that we do not know the true value of this parameter used in the simulation. In the process noise case the estimated value is -.3205. We know it is not the true $Z\alpha$, but is some value close to the true value. What additional information does a standard deviation of .0025 give us? If the Chebyshev inequality is used, we can be 75% certain that the true parameter lies within $\pm .3205 \pm .0050$, i.e., -.3155 and -.3255.

3.2.2 Lateral Motions

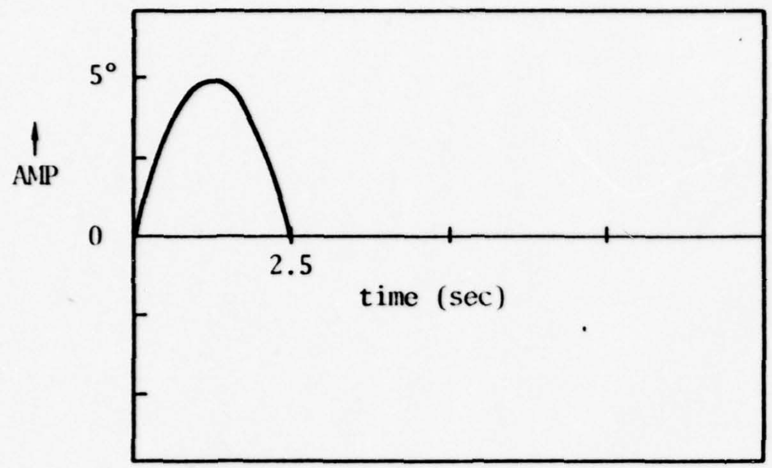
A rudder input followed by an aileron input of Figure 3.4 are used to excite the lateral motions of DC-8 aircraft. The parameter values used in the simulation are shown in Table 3.2. Only five parameters, N_r, N_β, Y_β, L_{δ_a} and N_{δ_r} are identified while the others are fixed at the least square values. The identified value and the least square starting value for the maximum likelihood program are also shown in Table 3.2.

The estimated value of N_β, Y_β, $L_{\delta a}$ and $N_{\delta r}$ is quite close to the true value. However, there is an error of 18.7% in the identified value of N_r. This does not necessarily mean that N_r is poorly identifiable. The error could occur because other parameters have been fixed at incorrect values in the maximum likelihood identification.

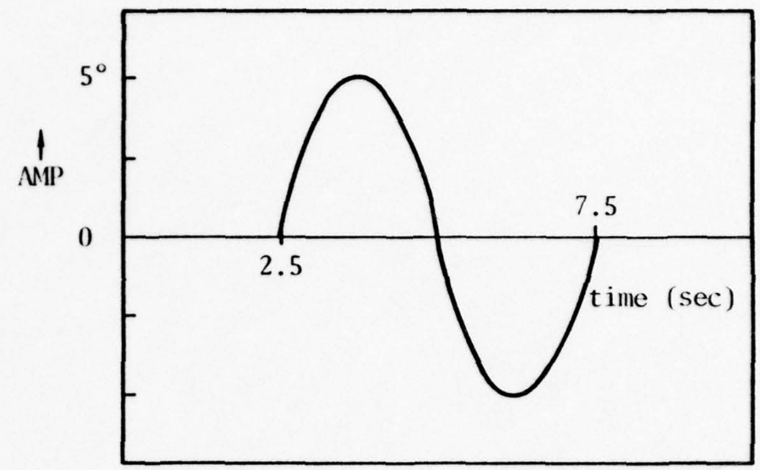
---

* The Chebyshev inequality states that a random variable will lie within $\pm x\sigma$ from the mean more than $(1 - \frac{1}{x^2})$ fraction of times. For example, a random variable will lie within $\pm 2\sigma$ more than 75% times. Note that this inequality is independent of the distribution function.

INPUT:



Rudder



Aileron

Figure 3.4 Rudder and Aileron Inputs for DC-8 Simulation

Table 3.2

PARAMETER IDENTIFICATION RESULTS FOR LATERAL CASE EXAMPLE

(SIMULATED DATA)

PARAMETER	TRUE VALUE	STARTING VALUE	IVA RESULT	ML RESULT	% ERROR
L_p	-2.029	-12.840	-2.105	*	
L_r	0.4697	- 2.700	.4407	*	
L	-2.092	1.628	-1.883	*	
N_p	-0.05828	- 7.770	-0.05089	*	
N_r	-0.2724	-17.540	-0.3107	-0.3233	18.7
N_β	1.655	1.360	1.716	1.693	2.3
Y_β	-0.1274	-20.1896	-0.1289	-0.1348	5.7
α	0.0	1.186	0.715	0.005	
L_{δ_a}	22.02	- 1.83	22.39	21.98	0.2
L_{δ_r}	0.4807	- 0.260	0.2576	*	
N_{δ_a}	-0.1709	1.0534	-0.2800	*	
N_{δ_r}	-1.3680	- 0.0071	-1.3784	-1.3116	4.1
Y_{δ_a}	0.0	0.2314	-1.274	*	
Y_{δ_r}	0.03630	0.0114	0.23232	*	

* Not identified

[All in units of ft., rad., sec.]

The predicted measurement time histories based on identified parameters are a good fit to the "true" measurements, as shown in Figure 3.5.

3.3 FLIGHT TEST DATA

The flight test data presented in this section is from a Navy swept-wing fighter at 30,000 feet altitude, 0.6 mach airspeed and 20° sweep angle. The wind tunnel estimates of aircraft stability and control coefficients are inaccurate.

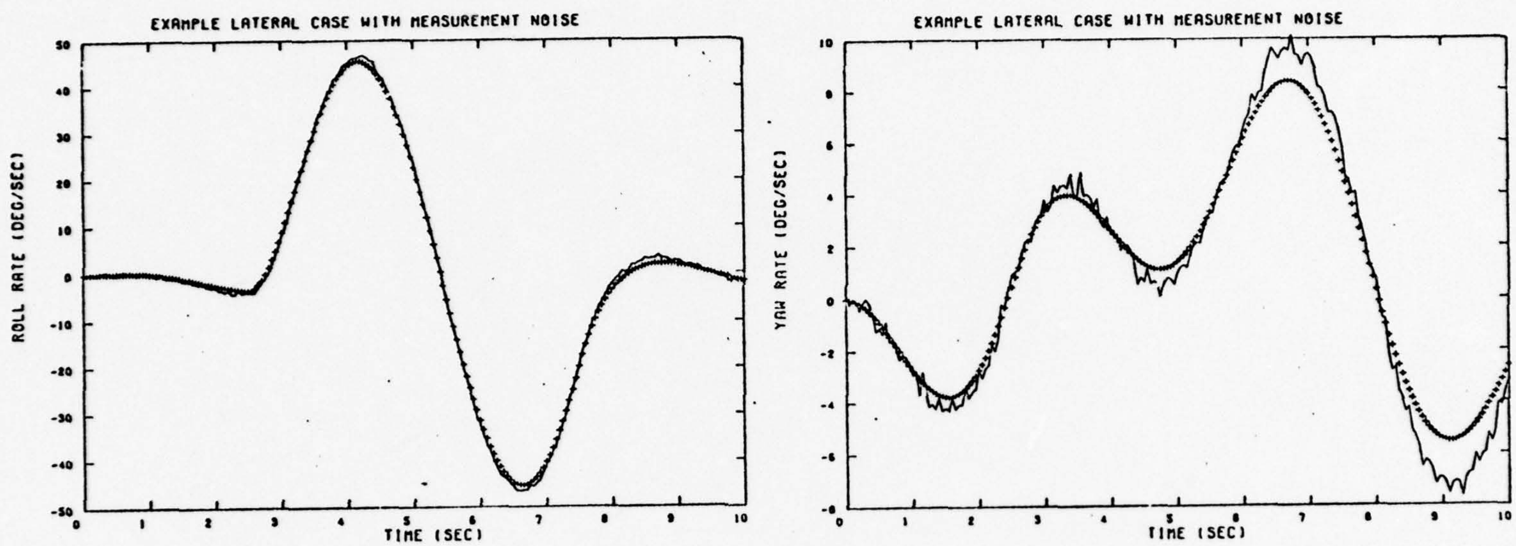


Figure 3.5 Lateral Simulation Data for DC-8

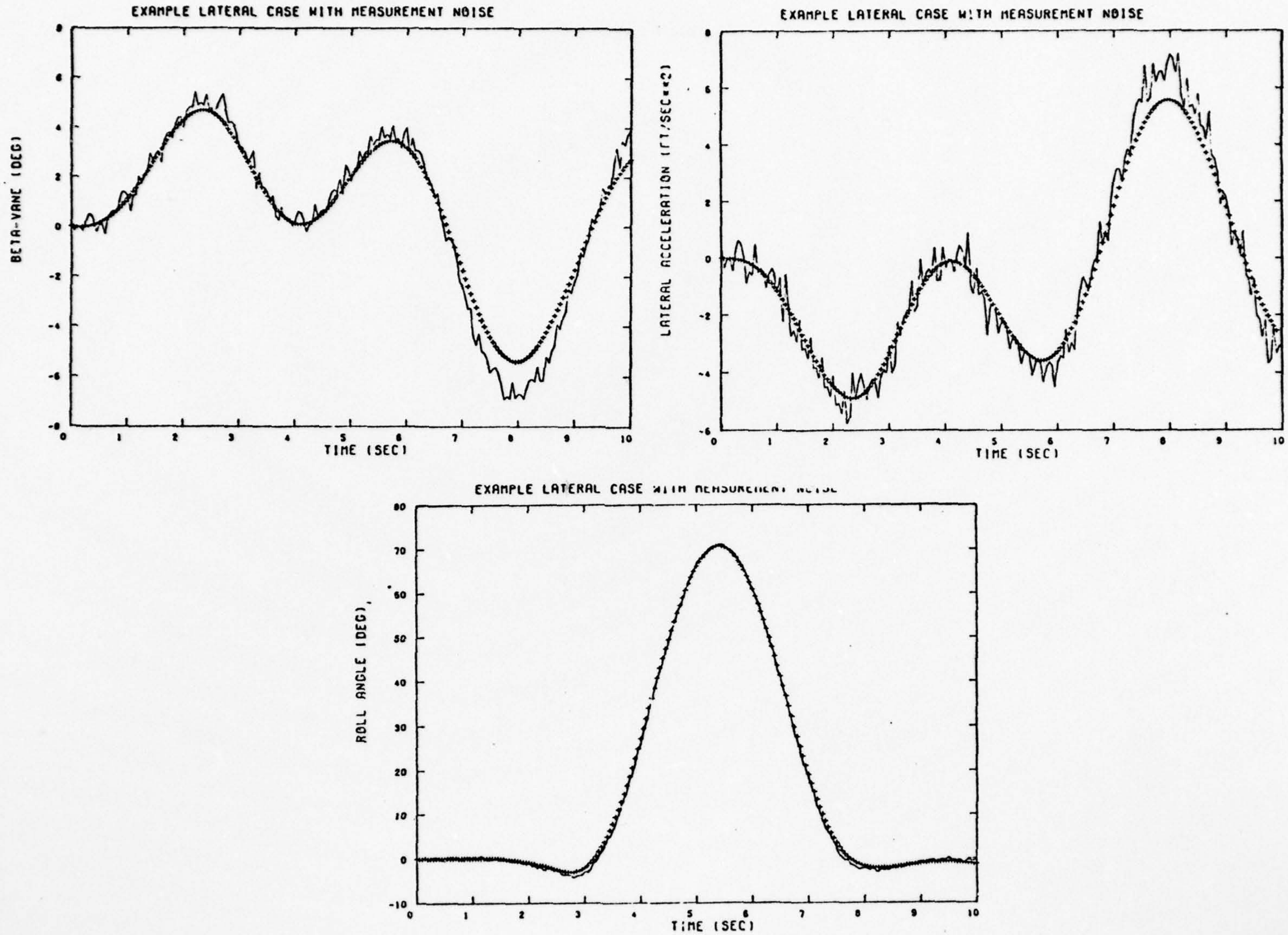


Figure 3.5 (cont'd) Lateral Simulation Data for DC-8

3.3.1 Longitudinal Motions Flight Data

The response data shown in Figure 3.7 is for a doublet input of Figure 3.6. The data is fairly noise free. "Spikes", arising from instrument recording dynamic response characteristics are evident.

Since no a priori parameter estimate of sufficient accuracy to start the maximum likelihood identification program were available, it was necessary to use a least squares approach to obtain starting values of the parameters. These values are shown in Table 3.3. The flight test data is processed as such to obtain the parameter estimates of Table 3.3 and predicted measurement time history of Figure 3.7. The estimates look reasonable and the fit to the actual measurement is quite good. The outliers which are caused by recording or instrumentation error can cause significant errors in parameter estimates. Therefore, the data is preprocessed before identification is performed. In this preprocessing the outliers are removed. In addition, one more parameter considered important is identified. The identified parameter values with data preprocessing is given in Table 3.3 and time history fits in Figure 3.8. There is a small change in most estimates. The one standard deviation estimation errors are enclosed in the brackets.

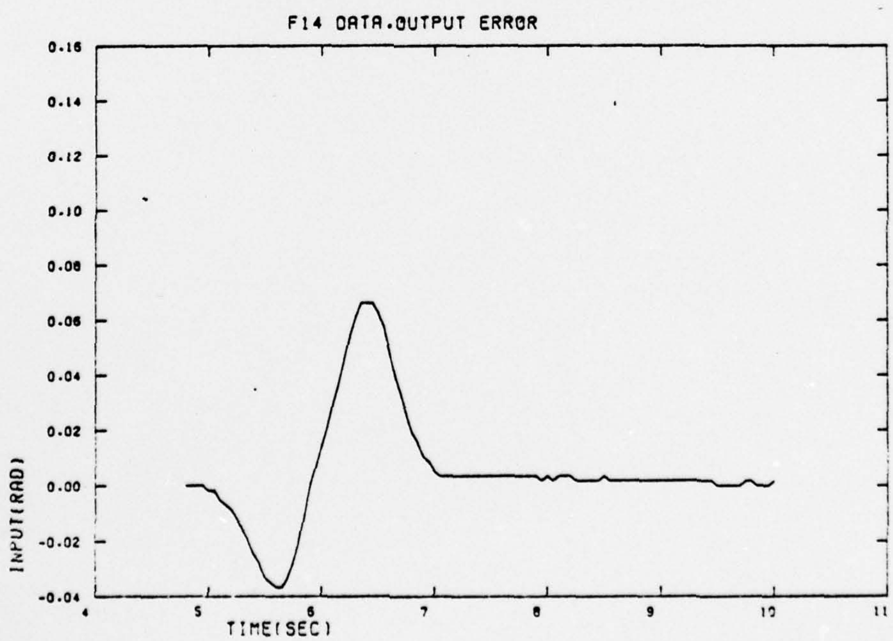


Figure 3.6 Elevator Input for Flight Test Data

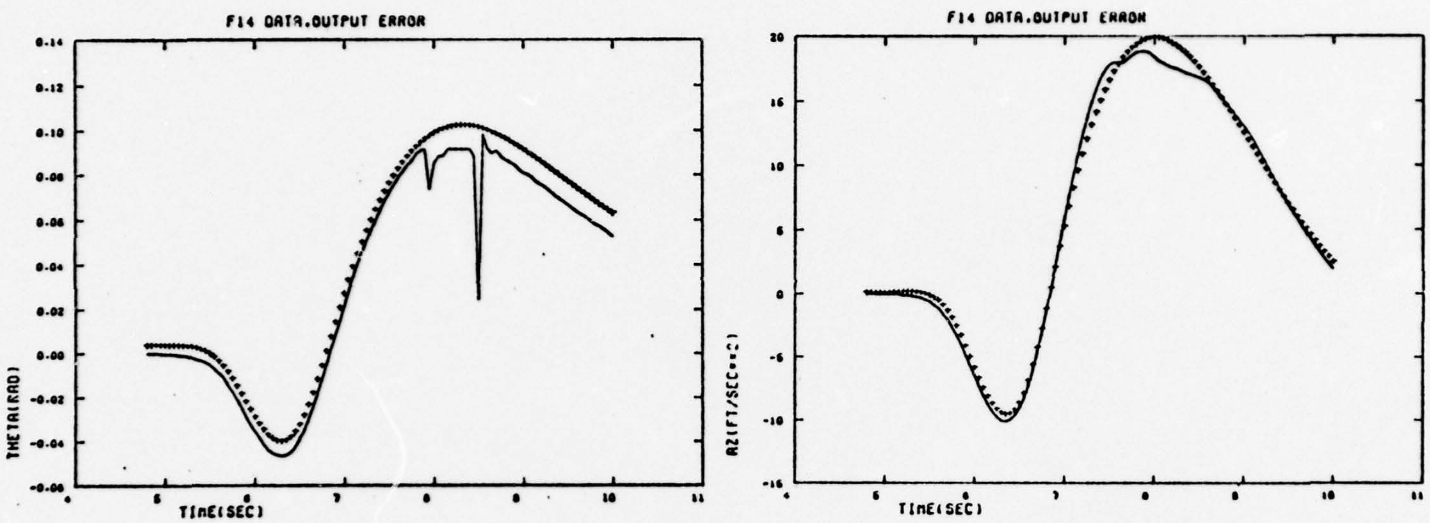


Figure 3.7 Flight Test Data With Output Error (Outliers Not Removed)

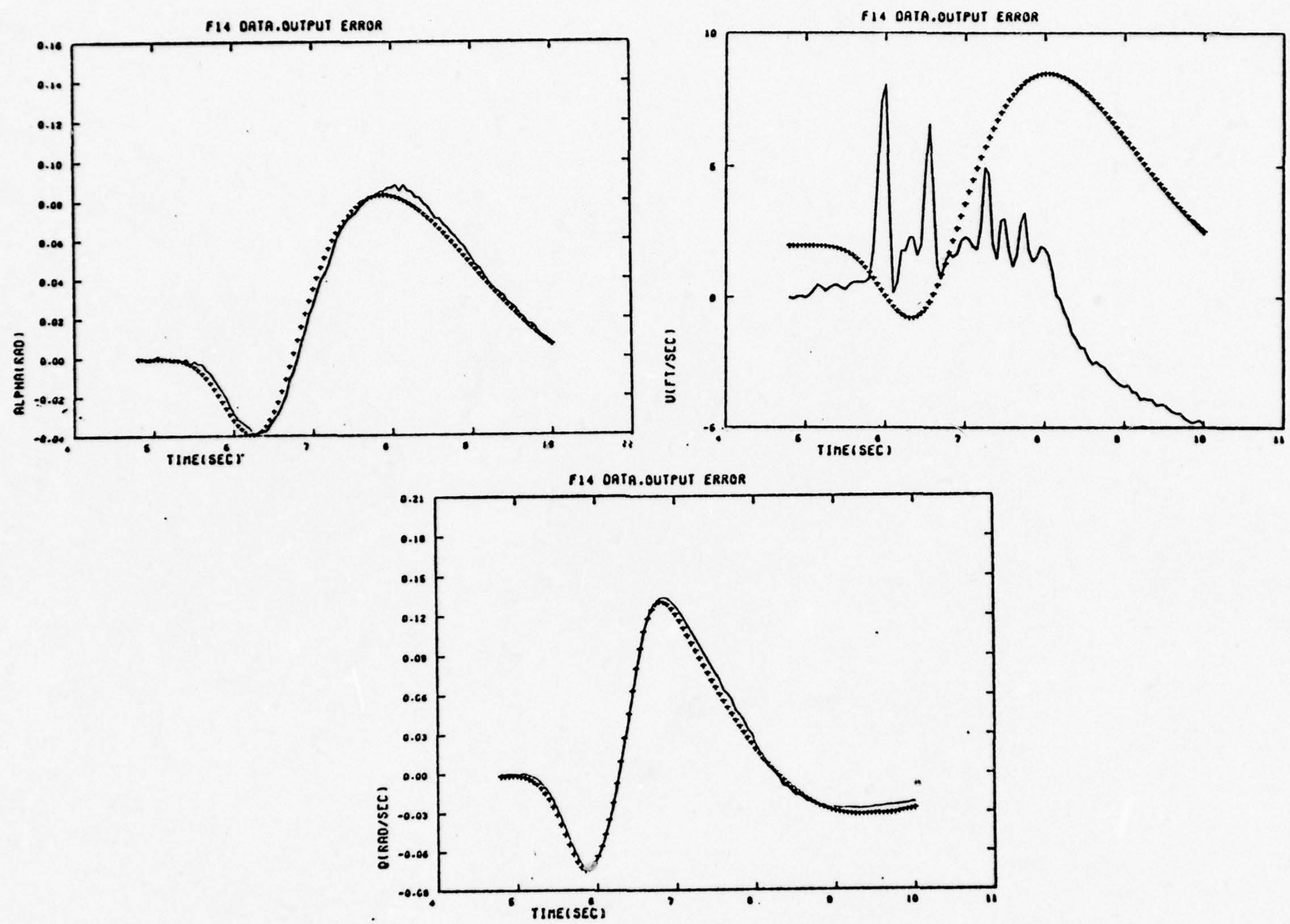


Figure 3.7 (cont'd) Flight Test Data With Output Error (Outliers Not Removed)

Table 3.3

F-14 RUN2

Altitude 30,000 ft.
Speed 362 knots
Sweepback 20°

DERIVATIVE	STARTING VALUE (LEAST SQUARES AND WIND TUNNEL)	ESTIMATED RAW DATA	VALUE (WITH 1σ BOUND) OUTLIERS REMOVED
Z_α	-0.38	-0.347 (.033)	-0.339 (.00271)
Z_u	-0.00066	*	*
Z_q	1.04	*	*
Y_o	0.0	*	*
X_α	285.0	*	*
X_u	-3.308	*	*
X_q	42.34	*	*
M_α	-0.755	-1.276 (.0789)	-1.285 (.00649)
M_u	0.00416	*	*
M_q	-0.600	-0.736 (.139)	-0.7068 (.011)
ω_c	0.50	*	*
Z_{δ_e}	-0.20	0.0378 (.218)	0.119 (.020)
X_{δ_e}	0.0	*	*
M_{δ_e}	-5.46	-5.357 (.452)	-5.338 (.033)
Q	0.0	*	*
k_α	1.0	*	*
$\frac{k_\alpha \ell_\alpha}{V}$	0.0	*	-0.0754 (.0062)
b_α	0.0	*	0.000124
b_u	0.0	*	-0.0934
b_q	0.0	*	0.00056
b_θ	0.0	*	0.000213
b_{a_z}	0.0	*	-0.05
σ_α	0.0075	0.0428	0.00159
σ_u	100.0	52.0	4.72
σ_q	0.007	0.0424	0.0038
σ_θ	0.015	0.0877	0.0021
σ_{a_z}	0.9	8.1	0.77

* Not identified

[All in units of ft., rad., sec.)

[Data Supplied by NATC TO SCI - April, 1974]

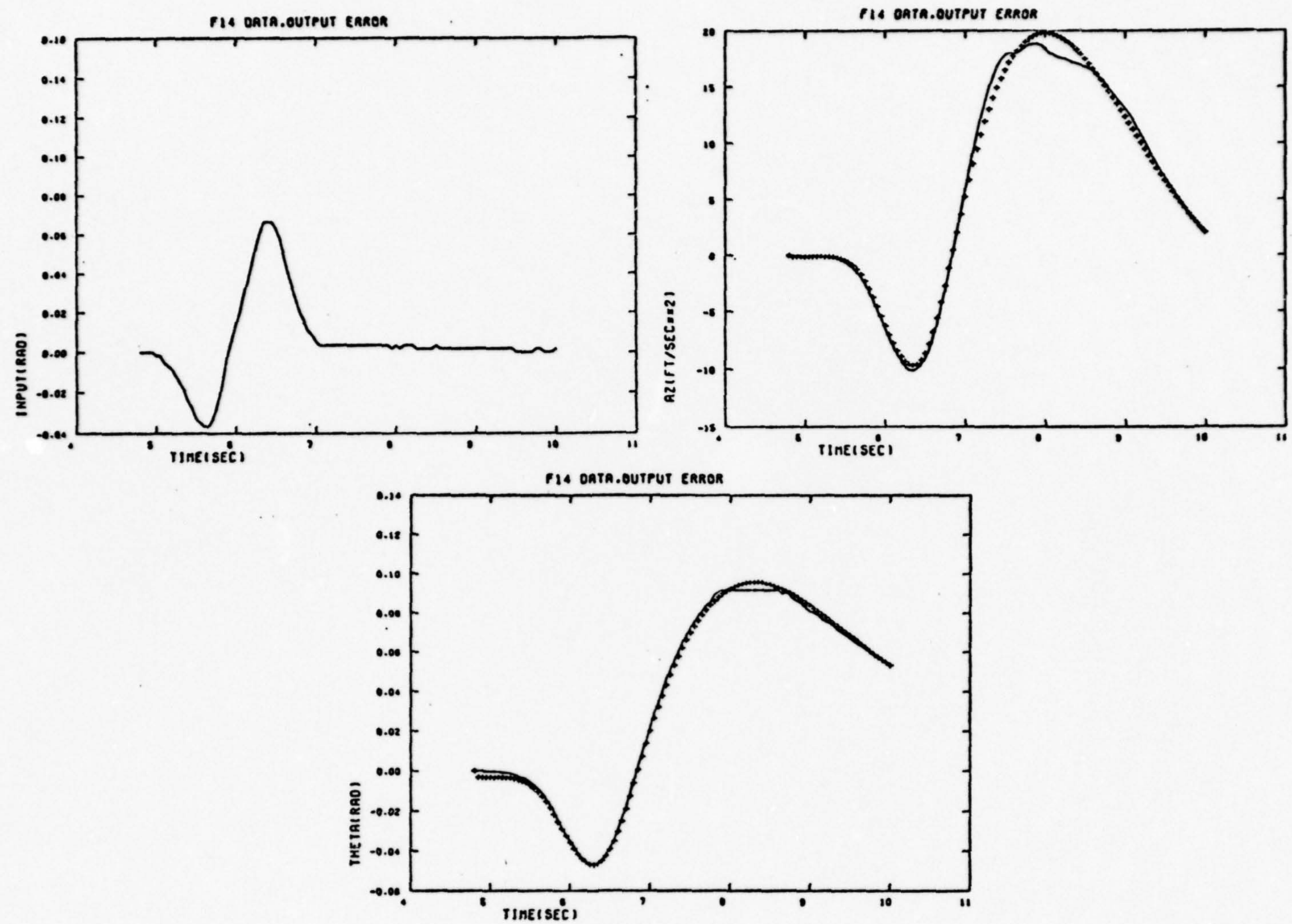


Figure 3.8 Flight Test Data (Outliers Removed)

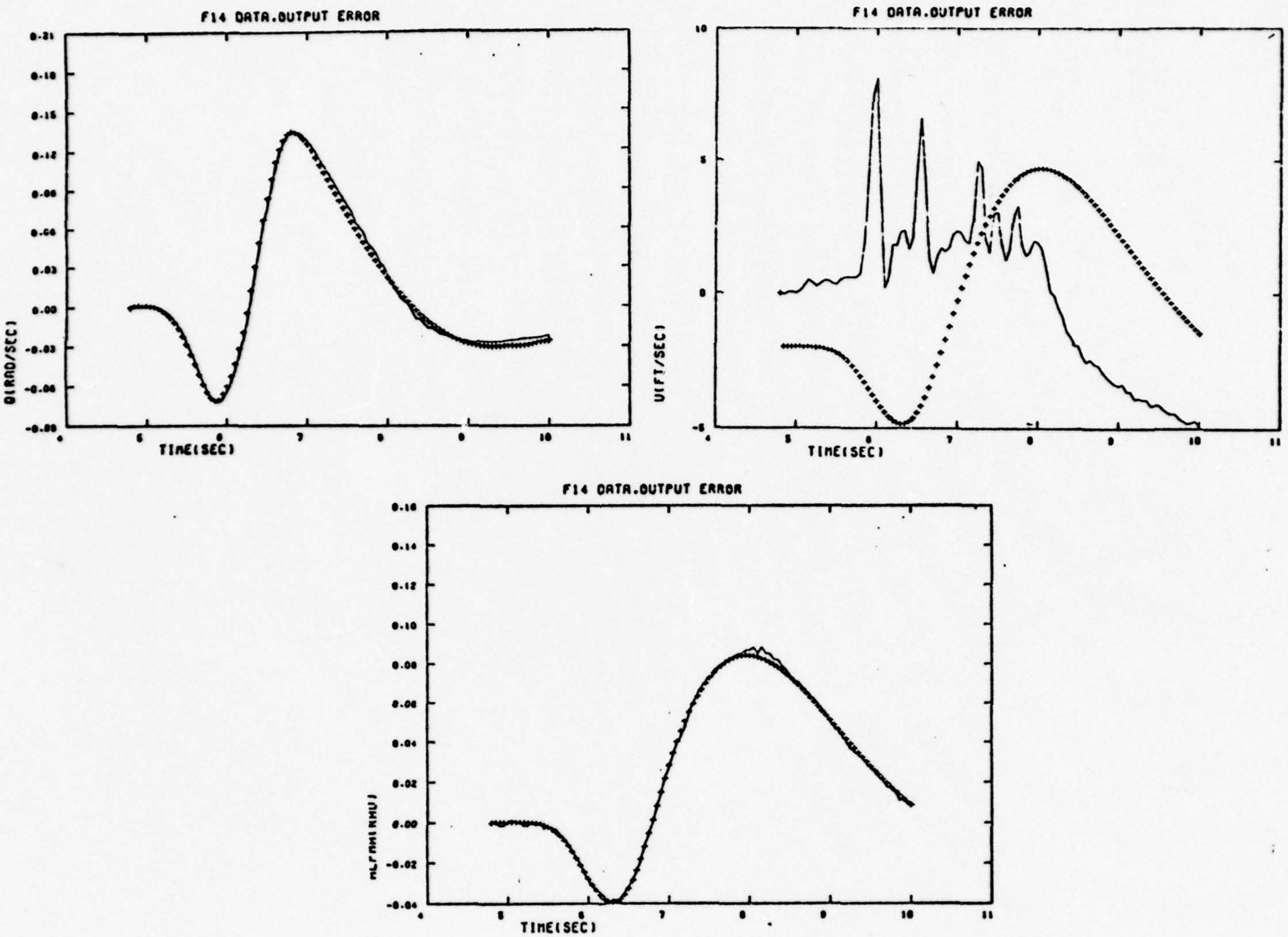


Figure 3.8 (cont'd) Flight Test Data (Outliers Removed)

Another response record resulting from a doublet input at approximately the same flight condition was obtained three months later. The identified value from previous run were used for startup. When the data was processed as such, it was immediately clear that the pitch rate measurement is bad. The raw data was rerun and the aircraft stability and control derivatives were identified without using the q measurement, but using the measurements of α, θ, u and nominal acceleration. The results are given in Table 3.4. Subsequently, the measurement of q was corrected (it had several sign errors) and also included in the identification to obtain parameter values shown in the last column of Table 3.4. The time history plots of the true and predicted measurements are given in Figure 3.9. Some parameters, e.g., $Z\alpha$ and $M\alpha$, are quite different from the last case.

3.3.2 Lateral Motions Flight Data

Lateral response data was obtained for rudder and aileron inputs of Figure 3.10. The data was processed, first, using the least squares approach to obtain the starting values for maximum likelihood identification. These values are used in the output error mode of the maximum likelihood method. This results in parameter estimates of Table 3.5 and time history plots of Figure 3.11. Next, the same data is processed, assuming lateral gusts. The reuslts are shown in Table 3.5 and Figure 3.12. The power spectral density of the process noise is .000347 sec^{-1}. Since the gust model break frequency is .5 sec, this gives a root mean square sideslip gust of

$$(\beta_g)RMS = \sqrt{\frac{.000347}{2x\ 5}} = .01863 \text{ rad} \qquad (3.1)$$

The aircraft speed is 362 knots. The lateral wind has a root mean square value of 7.68 ft sec^{-1}.

3.3.3 Verification of Parameter Estimates Obtained from Flight Data

The following steps are used to verify the validity of parameter estimates obtained from flight data.

Table 3.4

RUN1 (DOUBLET INPUT)

Altitude 30,000 ft.
Speed 360 knots
Sweepback 20°

PARAMETER	PARAMETER VALUE (WITH 1σ ERROR BOUND IN BRACKETS)			
	STARTING VALUE (FROM PREVIOUS RUN)	q FROM DATA (INCORRECT)	q NOT USED	q CORRECTED AND USED
Z_α	-0.3394	0.545	-0.4268 (.0097)	-0.4089 (.0059)
Z_u	-0.00066	*	*	*
Z_q	1.04	*	*	*
γ_o	0.0	*	*	*
X_α	285.0	*	*	*
X_u	-3.3	*	*	*
X_q	42.3	*	*	*
M_α	-1.28	-4.59	-2.002 (.018)	-1.88 (.0078)
M_u	0.0042	*	*	*
M_q	-0.74	-2.73	-0.4907 (.024)	-0.6726 (.0089)
ω_c	0.5	*	*	*
Z_{δ_e}	0.119	0.639	0.0856 (.014)	-0.0147 (.0136)
X_{δ_e}	0.0	*		
M_{δ_e}	5.34	0.416	4.154 (.093)	5.30 (.033)
Q	0.0	*	*	*
k_α	1.0	*	*	0.881 (.0089)
$\frac{k_\alpha l_\alpha}{v}$	0.0	-0.08	-0.0594 (.0066)	*
b_α	0.0	-0.0328	0.00609	-0.00071
b_u	0.0	-2.06	0.344	-0.147
b_q	0.0	-0.0104	0.00164	-0.000877
b_θ	0.0	-0.0378	0.0112	-0.00505
b_{a_z}	0.0	-0.81	0.922	-0.528
σ_α	0.0075	0.034	0.0072	0.00262
σ_u	100.0	2.7	4.05	1.35
σ_q	0.007	0.023	0.054	0.0034
σ_θ	0.015	0.063	0.0131	0.0100
σ_{a_z}	0.9	10.0	2.6	2.1

* Not identified

[All in units of ft., rad., sec.)

[Data Supplied by NATC TO SCI - July, 1974]

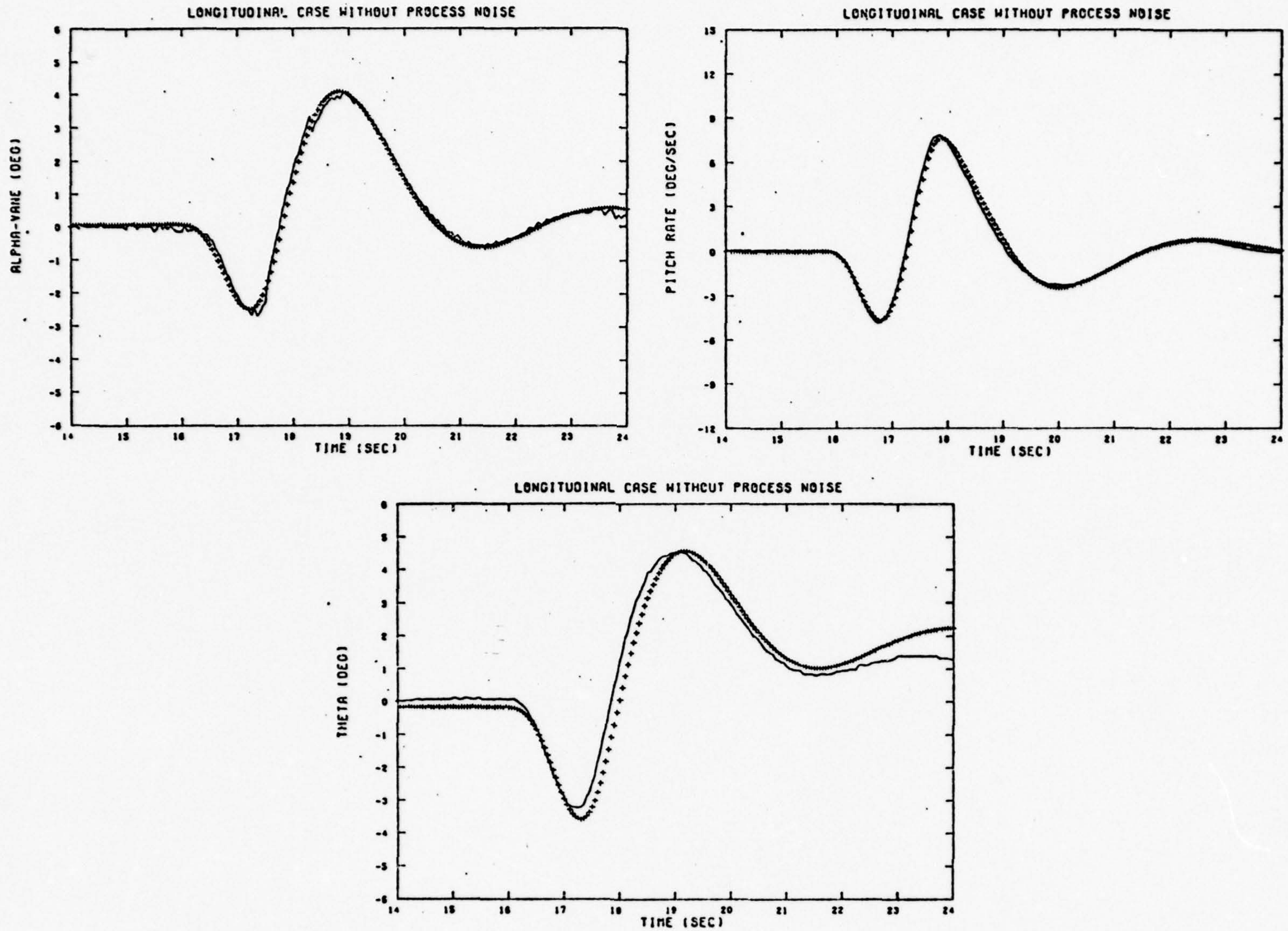


Figure 3.9 Longitudinal Flight Data Without Process Noise (Doublet Input)

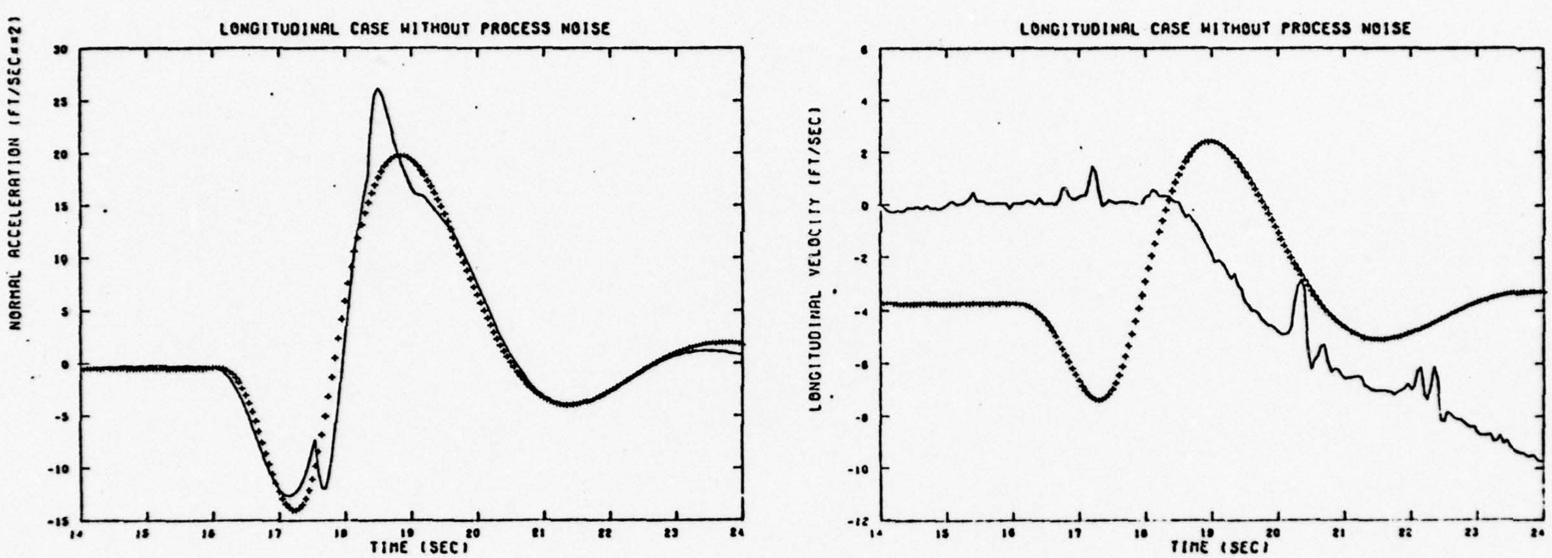


Figure 3.9 (cont'd) Longitudinal Flight Data Without Process Noise (Doublet Input)

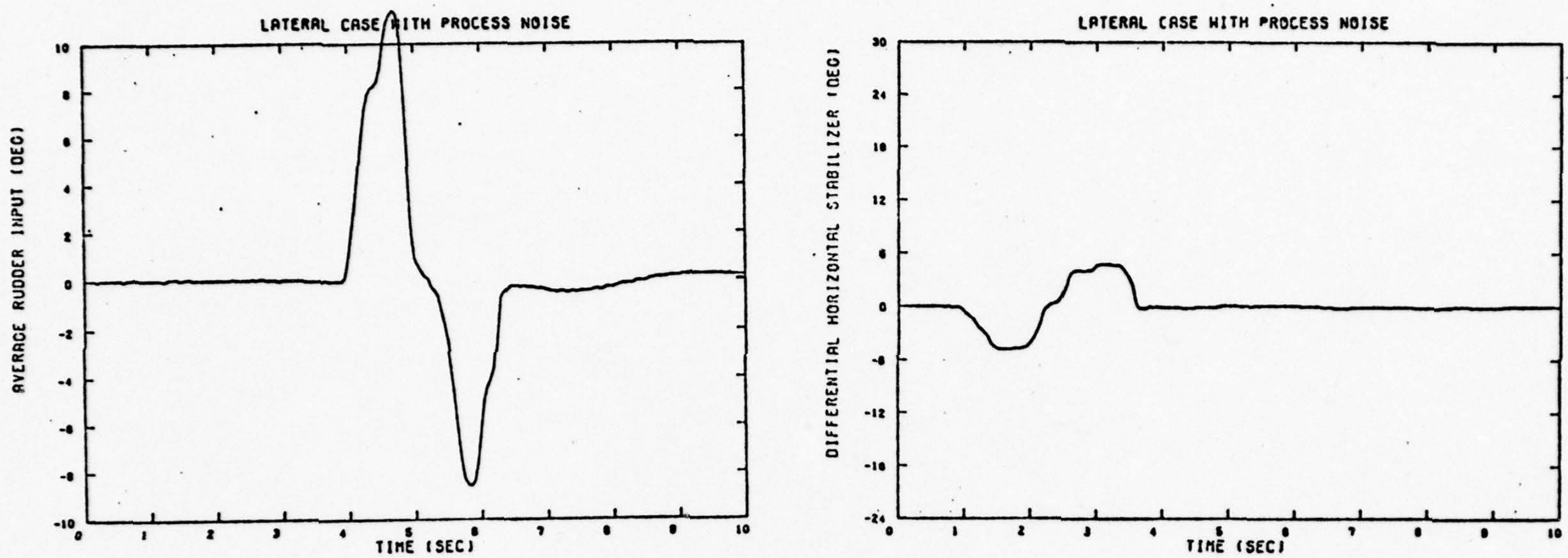


Figure 3.10 Rudder and Aileron Inputs for Lateral Excitation

Table 3.5

LATERAL MOTIONS (RUN8)*

Altitude 30,000 ft.
Speed 360 knots
Sweepback 20°

PARAMETER	STARTING VALUE (LEAST SQUARES)	MAXIMUM LIKELIHOOD ESTIMATE (WITH 1σ BOUND)	
		NO PROCESS NOISE	PROCESS NOISE
L_p	-1.975	-2.004 (.0235)	-1.948 (.0205)
L_r	4.877	2.686 (.0645)	2.369 (.0541)
L_β	-8.29	-8.500 (.0982)	-6.952 (.0846)
N_p	-0.004	*	*
N_r	-0.62	*	*
N_β	1.96	*	*
Y_β	-0.05	-0.0649 (.00190)	-0.0819 (.00121)
α_o	0.0	*	*
θ_o	0.0	*	*
L_{δ_a}	10.29	*	*
N_{δ_a}	0.24	*	*
Y_{δ_a}	0.0	*	*
L_{δ_r}	-2.828	-3.337 (.0865)	-1.446 (.0661)
N_{δ_r}	1.55	*	*
Y_{δ_r}	0.42	-0.015 (.0018)	-0.0150 (.00121)
K_β	1.0	*	*
$K_\beta \ell_\beta / V$	0.03	*	*
ω_c	0.5	*	*
Q	0.0	*	0.000347 (.000026)
b_p	-	-0.000209	-0.00264
b_r	-	-0.000126	0.000318
b_β	-	-0.0000590	0.000636
b_ϕ	-	-0.00230	-0.00516
b_{a_y}	-	0.00568	-0.0629
σ_p	0.005	0.0405	0.0246
σ_γ	0.005	0.0196	0.0069
σ_β	0.005	0.0190	0.00487
σ_ϕ	0.005	0.0192	0.0245
σ_{a_y}	0.5	1.57	0.85

* Not identified

[All in units of ft., rad., sec.)

[Data Supplied by NATC TO SCI - July, 1974]

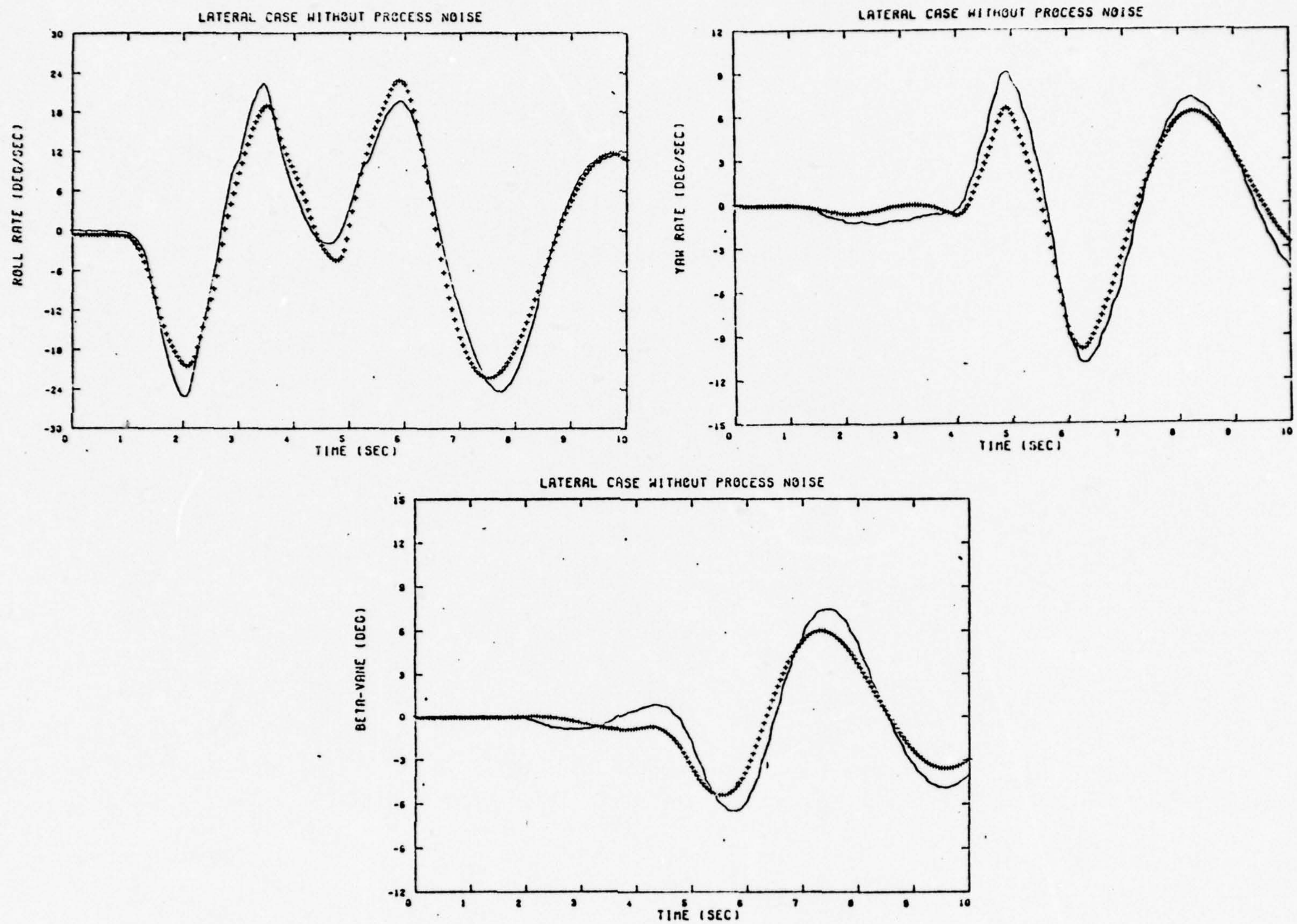


Figure 3.11 Measured and Predicted Aircraft Responses in the Lateral Direction (Output Error)

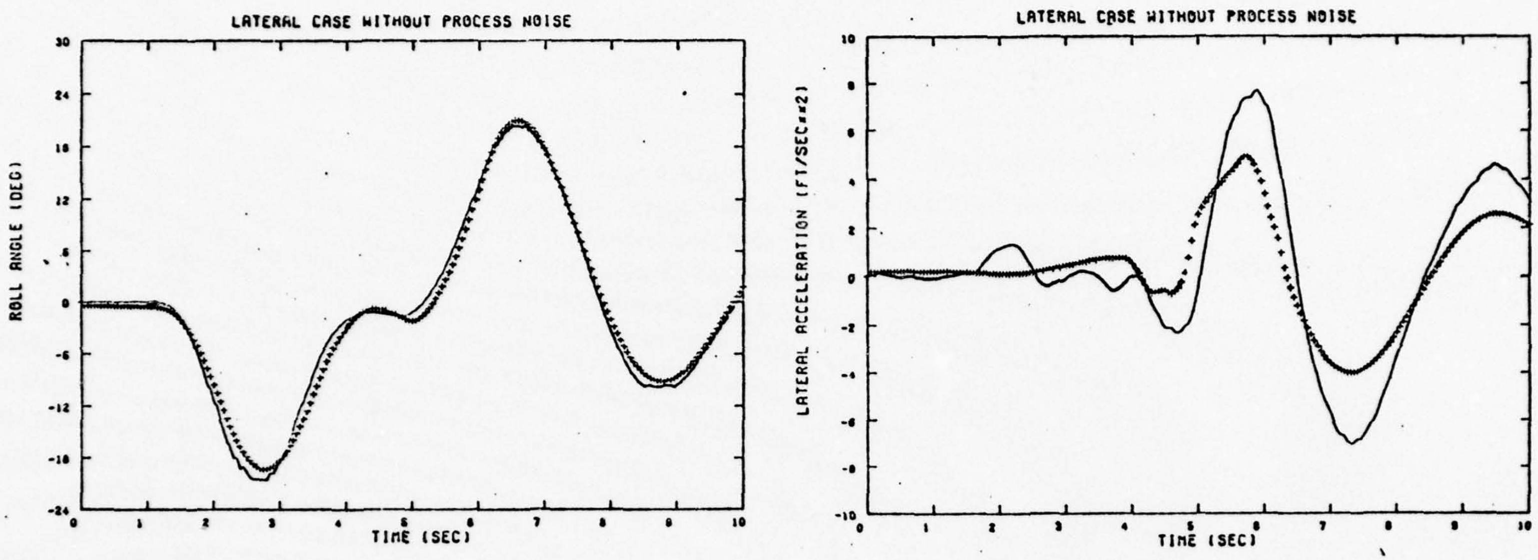


Figure 3.11 (cont'd) Measured and Predicted Aircraft Responses in the Lateral Direction (Output Error)

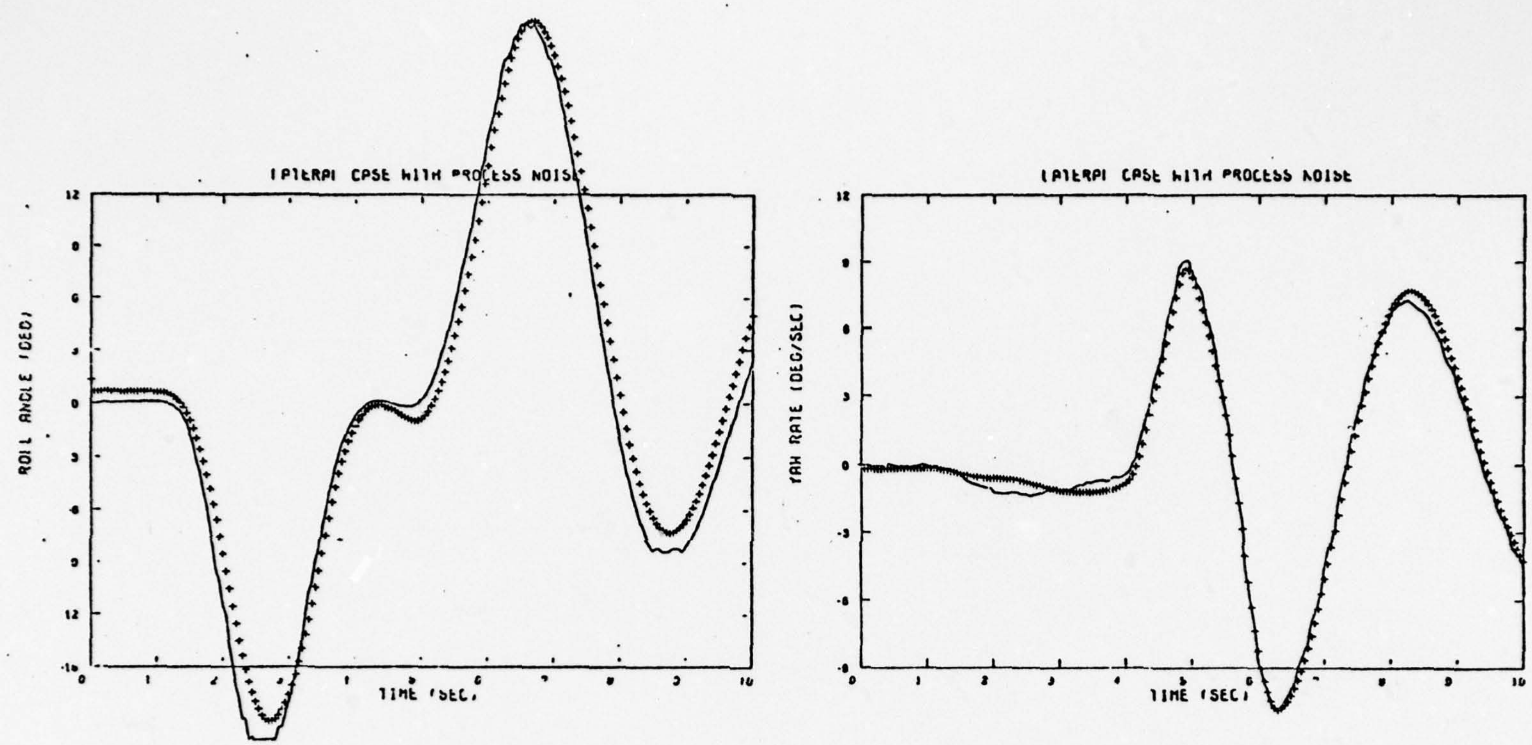


Figure 3.12 Measured and Predicted Aircraft Responses in the Lateral Direction (With Process Noise)

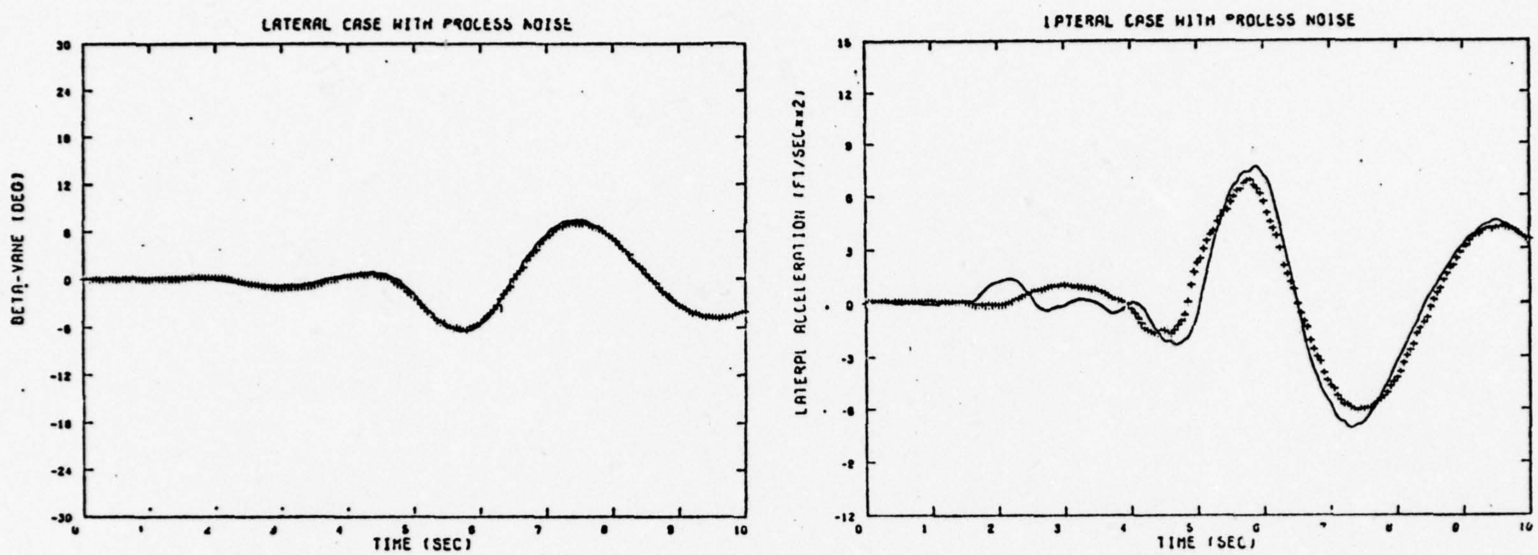


Figure 3.12 (cont'd) Measured and Predicted Aircraft Responses in the Lateral Direction (With Process Noise)

(a) The time history matches between the actual measurements and measurements predicted from estimated parameter values are good.

(b) The parameters are used to predict the frequency and damping of various modes. These are compared with values read from the response of different inputs. See Tables 3.6 and 3.7 for longitudinal and lateral motions, respectively.

(c) Response data from different inputs, but at the same flight condition is processed. The identified parameter values from different runs are close to each other; see Table 3.8 for the lateral case.

(d) Identified parameter value from response to one input are used to predict response to another input at the same flight condition. Again, excellent results were obtained.

TABLE 3.6

COMPARISON OF LONGITUDINAL SCIDNT
AND FLIGHT TEST RESULTS

DATA SOURCE	Run 1 (1)		Run 9 (2)		Run 12 (2)		Run 3 (1)
	ζ_{sp}	ω_{sp}	ζ_{sp}	ω_{sp}	ζ_{sp}	ω_{sp}	$P_{PHUGOID}$
SCIDNT	0.36	1.48	0.42	2.42	Not Used		86
FLIGHT TEST	0.38	1.36	0.37	2.37	0.3	2.51	94

NOTE: (1) 0.6 M, 30,000 ft, Load F, CR
(2) 0.6 M, 15,000 ft, Load F, CR

TABLE 3.7

COMPARISON OF LATERAL-DIRECTIONAL SCIDNT
AND FLIGHT TEST RESULTS

DATA SOURCE	$1/\tau_s$ (sec^{-1})	$1/\tau_R$ (sec^{-1})	ζ_d	ω_{n_d} (rad sec^{-1})
SCIDNT	0.0002	2.67	0.13	1.45
FLIGHT TEST	≈ 0	≈ 2.8	0.12	1.43

NOTE: (1) 0.6 m, 30,000 ft, CR, Load F

TABLE 3.8

PARAMETER ESTIMATES FROM DIFFERENT RUNS

(a) Stability Derivatives

RUN	L_p	L_r	L_β	N_p	N_r	N_β	Y_β	INPUTS
6B	-2.77	-	-	-	-	-	- .0825	Aileron Doublet
7E	-1.755	2.131	-6.157	.5976	- .052	3.90	.08062	Rudder Doublet
8B	-2.497	1.908	-8.17	- .004	- .376	1.94	- .0825	Aileron Rudder Doublet
8D	-2.62 (.04)	1.59 (.08)	-8.76 (.159)	- .0049 (.002)	- .353 (.0066)	1.959 (.013)	- .082 (.001)	Aileron Rudder Doublet

(b) Control Derivatives

RUN	L_{δ_a}	L_{δ_r}	N_{δ_a}	N_{δ_r}	Y_{δ_r}	Y_{δ_a}	INPUTS
6B	14.7	-	-	-	-	-	Aileron Doublet
7E	10.28	- .367	.24	1.51	- .0267	0.0	Rudder Doublet
8B	+13.03	-2.58	.24	1.55	- .0116	0.0	Aileron Rudder Doublet
8D	13.8 (.185)	-2.66 (.061)	.24 -	1.45 (.017)	- .0267 -	0.0	Aileron Rudder Doublet

IV. CONCLUSIONS

This report outlined the theoretical and statistical background of the maximum likelihood method for parameter estimation and showed its usefulness in identifying aircraft stability and control coefficients from flight response data. Several important conclusions can be drawn regarding the application of maximum likelihood method and, in particular, the computer program SCIDNT for reducing flight test data.

It is possible to obtain meaningful parameter estimates from well conducted flight tests using the maximum likelihood approach. These estimates are consistent between runs, can predict frequency and damping of aircraft natural modes and aircraft response to similar, but not same, inputs at the same flight condition (Section 3.3.3).

Both the output error and the process noise options of the maximum likelihood are important. While in most runs, the output error technique gives accurate parameter estimates and time history matches, in some runs significant improvement is achieved with the inclusion of process noise in the model. See Sections 3.2.1 and 3.3.2.

Because spurious signals are often present in flight data, some data preprocessing will usually be required. It is best, therefore, to make a visual examination of the recorded data before using any identification algorithm. The data should be checked for sign inconsistencies, outliers and other obvious errors. The dimensions of various signals could be checked for compatibility. The data preprocessor could vary in complexity (Section 3.3).

Input signals, used to excite the aircraft response, should be chosen with care. The set of parameters which can be identified and the accuracy with which they can be identified depends on the input (compare runs 6 and 8 in Table 3.8).

In summary, the computer implementation SCIDNT of the maximum likelihood parameter identification approach should prove useful in reducing flight test data to determine aircraft stability and control derivatives on a routine basis. It is important to understand, however, that the flight test must be planned with the objectives in mind that the response data must be carefully collected and checked out for successful evaluation of meaningful and useful aircraft stability and control coefficients.

APPENDIX A
MAXIMUM LIKELIHOOD IDENTIFICATION OF PARAMETERS

A.1 GENERAL NONLINEAR SYSTEMS

Consider the general nonlinear aircraft equations of motion

$$\dot{x} = f(x,u,\theta,t) + \Gamma(\theta,t)w \qquad 0 \leq t \leq T \tag{A.1}$$

$$E(x(0)) = x_o(\theta)$$

$$E\{(x(0) - x_o(\theta))(x(0) - x_o(\theta))^T\} = P_o(\theta) \tag{A.2}$$

where

$x(t)$ is $n \times 1$ state vector

$u(t)$ is $\ell \times 1$ input vector

θ is $p \times 1$ vector of unknown parameters

Γ is $n \times q$ process noise distribution matrix

$w(t)$ is $q \times 1$ random process noise vector

Sets of m measurements $y(t_k)$ are taken at discrete times t_k

$$y(t_k) = h(x(t_k),u(t_k),\theta,t_k) + v(t_k)$$

$$k = 1,2,3,\ldots,N \tag{A.3}$$

$w(t)$ and $v(t_k)$ are Gaussian random noises with the following properties

$$E(w(t)) = 0 \qquad E(v(t_k)) = 0 \qquad E(w(t)\ v^T(t_k)) = 0$$

$$E(w(t)\ w^T(\tau)) = Q(\theta,t)\ \delta(t-\tau),\ E(v(t_j)\ v^T(t_k)) = R(\theta,t_j)\delta_{jk} \tag{A.4}$$

The unknown parameters are supposed to occur in the functions f and h and in matrices Γ, Q, R, P_o and x_o. In the analysis to follow, the model and the functional form of f and h is assumed known correctly.

The set of observations $y(t_1), y(t_2), \ldots, y(t_N)$ constitutes the outcome Z in this case. The likelihood function for θ, which has the same form as the probability of the outcome z for a certain value of parameters θ, is given by

$$\begin{aligned} \mathscr{L}(\theta|z) &\equiv p(z|\theta) \\ &= p(y(t_1),\ y(t_2), \ldots, y(t_N)|\theta) \\ &= p(Y_N|\theta) \end{aligned} \tag{A.5}$$

where

$$Y_k = \{y(t_1), \ldots, y(t_k)\},\ k = 1,2,\ldots,N$$

$$\begin{aligned} p(Y_N|\theta) &= p(y(t_N)|Y_{N-1},\theta)\cdot p(Y_{N-1}|\theta) \\ &= p(y(t_N)|Y_{N-1},\theta)\ p(y(t_{N-1})|Y_{N-2},\theta)\ p(Y_{N-2}|\theta) \\ &= \prod_{i=1}^{N} p(y(t_i)|Y_{i-1},\theta) \end{aligned} \tag{A.6}$$

The log-likelihood function is

$$\log(\mathscr{L}(\theta|z)) = \sum_{i=1}^{N} \log\{p(y(t_i)|Y_{i-1},\theta)\} + \text{constant} \tag{A.7}$$

To find the probability distribution of $y(t_i)$ given Y_{i-1} and θ, the mean value and covariance are determined first.

$$E(y(t_i)|Y_{i-1},\theta) \triangleq \hat{y}(i/i-1) \tag{A.8}$$

The expected value or the mean is the best possible estimate of measurements at a point given the measurements until the previous point.

$$\begin{aligned} \mathrm{cov}(y(t_i)|Y_{i-1},\theta) &= E\{(y(t_i)-\hat{y}(i/i-1))(y(t_i)-\hat{y}(i/i-1))^T\} \\ &\triangleq E\{\nu(i)\ \nu(i)^T\} \\ &\triangleq B(i) \end{aligned} \qquad \text{(A.9)}$$

$\nu(i)$ are the innovations at point i and B(i) is the innovations covariance. Since

$$y(t_i) - E(y(t_i)|Y_{i-1},\theta) = \nu(i) \qquad \text{(A.10)}$$

it follows that $y(t_i)$ given Y_{i-1} and θ have the same distribution as $\nu(i)$. Kailath [44] has shown that as the sampling rate is increased, the innovations $\nu(i)$ tend towards having a Gaussian density. Assuming a sufficiently high sampling rate, the distribution of $\nu(i)$ and, therefore, $y(t_i)$ given Y_{i-1} and θ is Gaussian, i.e.,

$$p(y(t_i)|Y_{i-1},\theta) \equiv \frac{\exp\left\{-\frac{1}{2}\ \nu(i)^T\ B^{-1}(i)\ \nu(i)\right\}}{(2\pi)^{m/2}|B(i)|^{1/2}} \qquad \text{(A.11)}$$

$$\log\{p(y(t_i)|Y_{i-1},\theta)\} = -\frac{1}{2}\ \nu(i)^T B^{-1}(i)\nu(i) - \frac{1}{2}\log|B(i)| + \text{constant} \qquad \text{(A.12)}$$

The log-likelihood function of Equation (A.7) can be written as

$$\log(\mathscr{L}(\theta|z)) = -\frac{1}{2}\sum_{i=1}^{N}\{\nu^T(i)\ B^{-1}(i)\ \nu(i) + \log|B(i)|\} \qquad \text{(A.13)}$$

An estimate of the unknown parameters is obtained by maximizing the likelihood function or the log-likelihood function from the feasible set of parameter values.

$$\hat{\theta} = \max_{\theta\varepsilon\Theta} \log(\mathscr{L}(\theta|z)) \tag{A.14}$$

$$= \max_{\theta\varepsilon\Theta}\left[-\frac{1}{2}\sum_{i=1}^{N}\{\nu^T(i)B^{-1}(i)\nu(i) + \log|B(i)|\}\right] \tag{A.15}$$

The log-likelihood function depends on the innovations and their covariance. To optimize the likelihood function, a way must be found for determining these quantities. Both innovations and their covariance are outputs of an extended Kalman filter. This required Kalman filter is developed now.

The extended Kalman filter is conventionally divided into two parts. In the first part, called the prediction equations, the state equations and state estimate covariance equations are propagated in time from one measurement point to the next. In the second part, called the measurement update equations, the measurements and associated measurement noise covariances are used to improve state estimates. The covariance matrix is also updated at this point to reflect the additional information obtained from the measurements.

<u>Prediction Equations</u>

The state prediction is done using the equations of motion (A.2). Starting at time t_{i-1} with current estimate $\hat{x}(i-1|i-1)$ of the state $x(t_i)$ and the covariance $P(i-1|i-1)$, the following equations are used to find the predicted state $\hat{x}(i|i-1)$ and the associated covariance $P(i|i-1)$; see Bryson and Ho [45].

$$\frac{d}{dt}\hat{x}(t|t_{i-1}) = f(\hat{x}(t|t_{i-1}), u(t), \theta, t) \tag{A.16}$$

$$\dot{P}(t|t_{i-1}) = F(t)\, P(t|t_{i-1}) + P(t|t_{i-1})\, F^T(t) + \Gamma Q \Gamma^T \tag{A.17}$$

$$t_{i-1} \leq t \leq t_i$$

The nxn matrix F is obtained by linearizing f about the best current estimate

$$F(t) = \frac{\partial f(\hat{x}(t|t_{i-1}), u(t), \theta, t)}{\partial \hat{x}(t|t_{i-1})} \tag{A.18}$$

Using (A.16) to (A.18), we can obtain

$$\hat{x}(t_i|t_{i-1}) \triangleq \hat{x}(i|i-1)$$

and $P(t_i|t_{i-1}) \triangleq P(i|i-1)$ (A.19)

Thereafter, the measurement update equations are used.

Measurement Update Equations

The covariance and state estimate are updated using the measurements. The necessary relations are derived by Bryson and Ho [45] and are presented here without proof. The innovation and its covariance are

$$\nu(i) = y(i) - h(\hat{x}(i|i-1),\ u(t_i),\ \theta,\ t_i) \tag{A.20}$$

$$B(i) = H(i)\ P(i|i-1)\ H^T(i) + R \tag{A.21}$$

where H is obtained by linearizing h,

$$H(i) = \frac{\partial h(\hat{x}(i|i-1), u(t_i; \theta, t_i)}{\partial \hat{x}(i|i-1)} \tag{A.22}$$

The Kalman gain and the state update equations are

$$K(i) = P(i|i-1)\ H^T(i)\ B^{-1}(i) \tag{A.23}$$

$$\hat{x}(i|i) = \hat{x}(i|i-1) + K(i)\ \nu(i) \tag{A.24}$$

Finally, $P(i|i)$, the covariance of error in updated state, is obtained by

$$P(i|i) = (I-K(i)\ H(i))\ P(i|i-1) \tag{A.25}$$

One can now return to the time update or prediction equations.

A.3 OPTIMIZATION PROCEDURE

Many possible numerical procedures can be used for this optimization problem. Modified Newton-Raphson [20] or Quasilinearization [21] have been found by experience to give quicker convergence than most procedures like the conjugate gradient or the Davison method. The modified Newton-Raphson is a second order gradient procedure requiring computation of first and second order partials of the log-likelihood function.

$$\frac{\partial \log\ (\mathscr{L}(\theta|z))}{\partial \theta_j} = -\sum_{i=1}^{N} \left\{ \nu^T(i) B^{-1}(i) \frac{\partial \nu(i)}{\partial \theta_j} - \frac{1}{2} \nu^T(i) B^{-1}(i) \frac{\partial B(i)}{\partial \theta_j} B^{-1}(i) \nu(i) + \frac{1}{2} \mathrm{Tr}\left(B^{-1}(i) \frac{\partial B(i)}{\partial \theta_j} \right) \right\} \tag{A.26}$$

Also

$$\frac{\partial^2 \log \left(\mathscr{L}(\theta|z)\right)}{\partial\theta_j\,\partial\theta_k} = -\sum_{i=1}^{N}\left\{\frac{\partial\nu^T(i)}{\partial\theta_k} B^{-1}(i)\frac{\partial\nu(i)}{\partial\theta_j} - \frac{\partial\nu^T(i)}{\partial\theta_k} B^{-1}(i)\frac{\partial B(i)}{\partial\theta_j} B^{-1}(i)\nu(i)\right.$$

$$- \frac{\partial\nu(i)}{\partial\theta_j} B^{-1}(i)\frac{\partial B(i)}{\partial\theta_k} B^{-1}(i)\nu(i)$$

$$+ \nu^T(i)B^{-1}(i)\frac{\partial B(i)}{\partial\theta_j} B^{-1}(i)\frac{\partial B(i)}{\partial\theta_k} B^{-1}(i)\nu(i)$$

$$- \frac{1}{2}\,\mathrm{Tr}\left[B^{-1}(i)\frac{\partial B(i)}{\partial\theta_j} B^{-1}(i)\frac{\partial B(i)}{\partial\theta_k}\right] + \frac{\partial^2\nu(i)}{\partial\theta_j\partial\theta_k} B^{-1}(i)\nu(i)$$

$$- \nu^T(i)B^{-1}(i)\frac{\partial^2 B(i)}{\partial\theta_j\partial\theta_k} B^{-1}(i)\nu(i)$$

$$\left. + \frac{1}{2}\,\mathrm{Tr}\left(B^{-1}(i)\frac{\partial^2 B(i)}{\partial\theta_j\partial\theta_k}\right)\right\} \qquad j,k = 1,2,\ldots,p \qquad \text{(A.27)}$$

The last three terms in the equation for second partial of the log-likelihood function involve second partials of innovation and its covariance. Those terms are usually dropped. So the second partial is approximated by

$$\frac{\partial^2 \log(\mathscr{L}(\theta|z))}{\partial\theta_j\,\partial\theta_k} \cong - \sum_{i=1}^{N} \left\{ \frac{\partial\nu^T(i)}{\partial\theta_j} B^{-1}(i) \frac{\partial\nu(i)}{\partial\theta_k} - \frac{\partial\nu^T(i)}{\partial\theta_k} B^{-1}(i) \frac{\partial B(i)}{\partial\theta_j} B^{-1}(i)\nu(i) \right.$$

$$- \frac{\partial\nu(i)}{\partial\theta_j} B^{-1}(i) \frac{\partial B(i)}{\partial\theta_k} B^{-1}(i)\ \nu(i)$$

$$+ \nu^T(i) B^{-1}(i) \frac{\partial B(i)}{\partial\theta_j} B^{-1}(i) \frac{\partial B(i)}{\partial\theta_k} B^{-1}(i)\nu(i)$$

$$\left. - \frac{1}{2} \operatorname{Tr}\left[B^{-1}(i) \frac{\partial B(i)}{\partial\theta_j} B^{-1}(i) \frac{\partial B(i)}{\partial\theta_k} \right] \right\} \tag{A.28}$$

The gradients of innovation and its covariance for parameter θ_j are (Equations (A.20) and (A.21))

$$\frac{\partial\nu(i)}{\partial\theta_j} = - \left.\frac{\partial h}{\partial x}\right|_{x=\hat{x}(i|i-1)} \frac{\partial\hat{x}(i|i-1)}{\partial\theta} - \frac{\partial h}{\partial\theta} \tag{A.29}$$

$$j = 1,2,\ldots,p\ ;\ i = 1,2,\ldots,N$$

$$\frac{\partial B(i)}{\partial\theta_j} = \frac{\partial H(i)}{\partial\theta_j} P(i|i-1)H^T(i) + H(i) \frac{\partial P(i|i-1)}{\partial\theta_j} H^T(i)$$

$$+ H(i)\ P(i|i-1) \frac{\partial H^T(i)}{\partial\theta_j} + \frac{\partial R}{\partial\theta_j} \tag{A.30}$$

$$j = 1,2,\ldots,p$$

$$i = 1,2,\ldots,N$$

Recursive equations can be obtained for gradients of the predicted state and its covariance. This is done in stages by using the prediction and measurement update equations of Kalman filter. Differentiating (A.16) to (A.18) with respect to θ_j

$$\frac{d}{dt}\;\frac{\partial\hat{x}(t|t_{i-1})}{\partial\theta_j} = \frac{\partial f(\hat{x}(t|t_{i-1}),u(t),t,\theta)}{\partial\theta_j} + \frac{\partial f(\hat{x}(t|t_{i-1}),u(t),t,\theta)}{\partial\hat{x}(t|t_{i-1})}$$

$$\times\;\frac{\partial\hat{x}(t|t_{i-1})}{\partial\theta_j}$$

$$\frac{\partial\hat{x}(0|0)}{\partial\theta_j} = \frac{\partial x_o(\theta)}{\partial\theta_j} \qquad \text{(A.31)}$$

$$\frac{d}{dt}\;\frac{\partial P(t|t_{i-1})}{\partial\theta_j} = \frac{\partial F(t)}{\partial\theta_j}\,P(t|t_{i-1}) + F(t)\,\frac{\partial P(t|t_{i-1})}{\partial\theta_j} + \frac{\partial P(t|t_{i-1})}{\partial\theta_j}\,F^T(t)$$

$$+\,P(t|t_{i-1})\,\frac{\partial F^T(t)}{\partial\theta_j} + \frac{\partial\Gamma}{\partial\theta_j}\,Q\Gamma^T + \Gamma\,\frac{\partial Q}{\partial\theta_j}\,\Gamma^T + \Gamma Q\,\frac{\partial\Gamma^T}{\partial\theta_j}$$

$$\frac{\partial P(0|0)}{\partial\theta_j} = \frac{\partial P_o(\theta)}{\partial\theta_j} \qquad t_{i-1} \le t \le t \qquad \text{(A.32)}$$

$$\frac{\partial F(t)}{\partial\theta_j} = \frac{\partial^2 f(\hat{x}(t|t_{i-1}),u(t),\theta,t)}{\partial\theta_j\;\partial\hat{x}(t|t_{i-1})} \qquad j = 1,2,\ldots,p \qquad \text{(A.33)}$$

The sensitivity functions are updated at measurement points by differentiating (A.22) to (A.25) with respect to θ_j.

$$\frac{\partial H(i)}{\partial\theta_j} = \frac{\partial^2 h(\hat{x}(i|i-1),u(t_i),\theta,t_i)}{\partial\theta_j\;\partial\hat{x}(i|i-1)} \qquad \text{(A.34)}$$

$$\frac{\partial K(i)}{\partial\theta_j} = \frac{\partial P(i|i-1)}{\partial\theta_j}\,H^T(i)\;B^{-1}(i) + P(i|i-1)\,\frac{\partial H^T(i)}{\partial\theta_j}\,B^{-1}(i)$$

$$-P(i|i-1)\;H^T(i)\;B^{-1}(i)\;\frac{\partial B(i)}{\partial\theta_j}\,B^{-1}(i) \qquad \text{(A.35)}$$

$$\frac{\partial \hat{x}(i|i)}{\partial \theta_j} = \frac{\partial \hat{x}(i|i-1)}{\partial \theta_j} + \frac{\partial K(i)}{\partial \theta_j} \nu(i) + K(i) \frac{\partial \nu(i)}{\partial \theta_j} \tag{A.36}$$

$$\frac{\partial P(i|i)}{\partial \theta_j} = (I-K(i)\ H(i)) \frac{\partial P(i|i-1)}{\partial \theta_j} - \frac{\partial K(i)}{\partial \theta_j} H(i)\ P(i|i-1)$$

$$- K(i) \frac{\partial H(i)}{\partial \theta_j} P(i|i-1) \tag{A.37}$$

$$j = 1,2,3,\ldots,p$$

The negative of the matrix of second partials of the log-likelihood function is called the information matrix M. The step size $\Delta\theta$ for parameter estimates is given by

$$\Delta\theta = M^{-1} \frac{\partial \log (\mathscr{L}(\theta|z)}{\partial \theta} \tag{A.38}$$

The information matrix provides a lower bound on parameter estimate covariances, i.e.,

$$E(\theta-\hat{\theta})(\theta-\hat{\theta})^T \geq M^{-1} \tag{A.39}$$

This is called the Cramer-Rao lower bound. [46,47] The maximum likelihood estimates approach this bound asymptotically.

A.4 LINEAR SYSTEMS

In a linear system, the functions f and h are defined as

$$f(x,u,\theta,t) \triangleq F(\theta,t)x + G(\theta,t)u$$

$$\text{and} \quad h(x,u,\theta,t) \triangleq H(\theta,t)x + D(\theta,t)u \tag{A.40}$$

In this case, it has been shown that if the model is correct, the innovations are white and have Gaussian density at the true values of the parameters. Therefore, the assumption of fast sampling rate is not necessary.

The basic algorithm is the same. However, some of the equations can now be simplified. The equivalence between F(t) and H(i) of (A.18) and (A.22) and $F(\theta,t)$ and $H(\theta,t_i)$ of (A.40) is obvious. Equations (A.16) and (A.20) now become

$$\frac{d}{dt}\hat{x}(t|t_{i-1}) = F(\theta,t)\,\hat{x}(t|t_{i-1}) + G(\theta,t)\,u(t) \tag{A.41}$$

and

$$\nu(i) = y(i) - H(\theta,t_i)\,\hat{x}(i|i-1) - D(\theta,t_i)\,u(t_i) \tag{A.42}$$

Equations (A.29) and (A.31) can be written as

$$\frac{\partial\nu(i)}{\partial\theta_j} = -H(\theta,t_i)\frac{\partial\hat{x}(i|i-1)}{\partial\theta_j} - \frac{\partial H(\theta,t_i)}{\partial\theta_j}\hat{x}(i|i-1) - \frac{\partial D(\theta,t_i)}{\partial\theta_j}u(t_i) \tag{A.43}$$

and

$$\frac{d}{dt}\frac{\partial\hat{x}(t|t_{i-1})}{\partial\theta_j} = F(\theta,t)\frac{\partial\hat{x}(t|t_{i-1})}{\partial\theta_j} + \frac{\partial F(\theta,t)}{\partial\theta_j}\hat{x}(t|t_{i-1}) + \frac{\partial G(\theta,t)}{\partial\theta_j}u(t) \tag{A.44}$$

All other equations remain the same.

There is considerable reduction in computation requirement for time-invariant linear system. In this case, matrices F, G, H, D, Γ, Q and R and

their derivatives with respect to parameters are constant. Gupta [48] has shown that the computation of state sensitivities can be reduced to many fewer equations. Similar reductions are possible in the computation of covariance sensitivities.

A.5 TIME INVARIANT LINEAR SYSTEMS IN STATISTICAL STEADY STATE

In many aircraft applications, the Kalman filter is in steady state for the duration of the experiment. This occurs when the Kalman filter is in operation for a sufficiently long time and the process and measurement noise covariances do not change. The Kalman gain and the innovations and the state covariances approach constant values. The time update and measurement update equations for the covariances are

$$\frac{d}{dt} P(t|t_{i-1}) = FP(t|t_{i-1}) + P(t|t_{i-1})F^T + \Gamma Q \Gamma^T \tag{A.45}$$

$$K = P(i|i-1)H^T B^{-1} \tag{A.46}$$

$$B = HP(i|i-1)H^T + R \tag{A.47}$$

$$P(i|i) = (I-KH)\, P(i|i-1) \tag{A.48}$$

By definition of the steady state

$$P(i-1|i-1) = P(i|i)$$

Therefore, from (A.45)

$$\begin{aligned} P(i|i-1) &= e^{F\Delta t} P(i-1|i-1)\, e^{F^T \Delta t} + \int_{t_{i-1}}^{t_i} e^{F(t_i-\tau)}\, \Gamma Q \Gamma^T e^{F^T (t_i-\tau)} d\tau \\ &= e^{F\Delta t} (I-KH)\, P(i|i-1) e^{F^T \Delta t} + \int_{t_{i-1}}^{t_i} e^{F(t_i-\tau)} \Gamma Q \Gamma^T\, e^{F^T (t_i-\tau)} d\tau \end{aligned} \tag{A.49}$$

$$\triangleq \varphi(\Delta t)\ (I-KH)\ P(i|i-1)\ \varphi(\Delta t) + Q' \tag{A.49}$$

$$= \varphi(\Delta t)\ (P(i|i-1) - KBK^T)\ \varphi(\Delta t) + Q'$$

Using (A.46), (A.47) and (A.49), we can solve for $P(i|i-1)$ and then find K and B. Also, it can be shown that

$$\frac{\partial P}{\partial \theta_j} = A_1 \frac{\partial P}{\partial \theta_j} A_1^T + A_2 - \varphi P A_3 P \varphi^T \tag{A.50}$$

$$\frac{\partial K}{\partial \theta_j} = (I-KH) \frac{\partial P}{\partial \theta_j} H^T (HPH^T + R)^{-1} + A_2 \tag{A.51}$$

$$\frac{\partial B}{\partial \theta_j} = \frac{\partial H}{\partial \theta_j} PH^T + H \frac{\partial P}{\partial \theta_j} H^T + HP \frac{\partial H^T}{\partial \theta_j} + \frac{\partial R}{\partial \theta_j} \tag{A.52}$$

where

$$A_1 = \varphi(I-KH)$$

$$A_2 = \frac{\partial \varphi}{\partial \theta_j} (I-KH)P\varphi^T + \varphi(I-KH)P \frac{\partial \varphi^T}{\partial \theta_j} - \varphi K \frac{\partial H}{\partial \theta_j} P\varphi^T + \frac{\partial Q'}{\partial \theta_j}$$

$$A_3 = \frac{\partial H^T}{\partial \theta_j} B^{-1}H + H^T B^{-1} \frac{\partial H}{\partial \theta_j} PH^T + HP \frac{\partial H^T}{\partial \theta_j} + \frac{\partial R}{\partial \theta_j} B^{-1}H \tag{A.53}$$

$$P \triangleq P(i|i-1)$$

Thus, it is possible to solve for $\frac{\partial P}{\partial \theta_j}$ using (A.50) and then find $\frac{\partial K}{\partial \theta_j}$ and $\frac{\partial B}{\partial \theta_j}$ from (A.51) and (A.52). (A.50) is a linear equation in $\frac{\partial P}{\partial \theta_j}$ and the coefficient of the unknown matrix does not depend on the parameter θ_j. Thus, the sensitivity of state covariance matrix can be determined very quickly for all parameters. Once the sensitivity of P, K and B for unknown parameters

is determined, only state sensitivity equations need to be updated. The computation of state sensitivity functions can be reduced to many fewer equations as pointed out in Section A.4.

An <u>approximation</u> suggested by Mehra [34] simplifies the problem further. The unknown parameters are defined to include elements in K and B matrices instead of Q and R. Optimizing the log-likelihood function for parameters in B gives

$$\hat{B} = \frac{1}{N} \sum_{i=1}^{N} \nu(i)\, \nu^T(i) \tag{A.54}$$

The gradient of the log-likelihood function with respect to other unknown parameters is

$$\frac{\partial \, \mathrm{Log}(\mathscr{L}(\theta|z))}{\partial \theta_j} = -\sum_{i=1}^{N} \nu^T(i)\, \hat{B}^{-1} \frac{\partial \nu(i)}{\partial \theta_j} \tag{A.55}$$

The sensitivity of innovations to parameters is determined using the following recursive equations

$$\frac{d}{dt}\hat{x}(t|t_{i-1}) = F\hat{x}(t|t_{i-1}) + Gu(t)$$

$$\frac{d}{dt}\frac{\partial \hat{x}(t|t_{i-1})}{\partial \theta_j} = \frac{\partial F}{\partial \theta_j}\hat{x}(t|t_{i-1}) + F\frac{\partial \hat{x}(t|t_{i-1})}{\partial \theta_j} + \frac{\partial G}{\partial \theta_j}u(t)$$

$$j = 1,2,\ldots,p \qquad t_{i-1} \le t \le t_i \tag{A.56}$$

$$\begin{aligned} \nu(i) &= y(i) - H\hat{x}(i|i-1) - Du(t_i) \\ \hat{x}(i|i) &= \hat{x}(i|i-1) + K\nu(i) \end{aligned} \tag{A.57}$$

$$\frac{\partial \nu(i)}{\partial \theta_j} = - \frac{\partial H}{\partial \theta_j} \hat{x}(i|i-1) - H \frac{\partial \hat{x}(i|i-1)}{\partial \theta_j} - \frac{\partial D}{\partial \theta_j} u(t_i)$$

$$\frac{\partial \hat{x}(i|i)}{\partial \theta_j} = \frac{\partial \hat{x}(i|i-1)}{\partial \theta_j} + \frac{\partial K}{\partial \theta_j} \nu(i) + K \frac{\partial \nu(i)}{\partial \theta_j} \tag{A.58}$$

$$j = 1,2,\ldots,p$$

Note that

$$\begin{aligned} \frac{\partial K}{\partial \theta_j} &= 0 \quad \text{if } \theta_j \text{ is not an element of K matrix} \\ &= I_{j'k'} \quad \text{if } \theta_j \triangleq K_{j'k'} \end{aligned} \tag{A.59}$$

where $I_{j'k'}$ is a matrix of all zeroes except a 1 at the j',k' position.

This approximation simplifies the optimization considerably. However, this usually leads to an overparameterized model. In other words, once K and B are determined, it is not possible to find any corresponding Γ, Q and R which have the desired (known a priori) structure. A good estimate of elements in Γ, Q and R matrices can be obtained by a least-squares type approach. The fit to the observed data is better than with true values of parameters but the parameter estimates do not have minimum variance. This approximation is not good in aircraft application where the structure of Q and R is known fairly well, but is excellent where there are many process noise sources and the characteristics of both Q and R are relatively unknown (e.g., economic systems).

A.6 MAXIMUM LIKELIHOOD WITH NO PROCESS OR MEASUREMENT NOISE

The maximum likelihood method can be simplified when either process noise or measurement noise are absent.

No Process Noise

If the process noise is zero and initial states are known perfectly, i.e., w(t) and P(0) are zero, the covariance of the error in the predicted state is also zero. It is clear from (A.23) that Kalman gains are zero. The innovations is the output error, i.e.,

$$\nu(i) = y(i) - h(x(t_i), u(t_i), \theta, t_i) \tag{A.60}$$

and the innovation covariance is (A.21)

$$B(i) = R \tag{A.61}$$

the log-likelihood function is,

$$\log\,(\mathscr{L}(\theta|z)) = -\frac{1}{2}\sum_{i=1}^{N} \nu^T(i)\, R^{-1}\nu(i) + \log|R| \tag{A.62}$$

which on optimizing for unknown parameters in R gives

$$\hat{R} = \frac{1}{N}\sum_{i=1}^{N} \nu(i)\ \nu^T(i) \tag{A.63}$$

The equality in (A.63) holds only for those elements of R which are not known a priori. For instance, even if R is known to be diagonal, the right hand side matrix will not be diagonal in general, but the off-diagonal terms should be ignored before they are equated to $\hat{R}$. Using (A.63) in (A.62)

$$\log(\mathscr{L}(\theta|z)) = -\frac{1}{2}\sum_{i=1}^{N} \nu^T(i)\ \hat{R}^{-1}\,\nu(i) + \text{constant} \tag{A.64}$$

The optimizing function is the same as that for the output error method except that the measurement noise covariance matrix is determined using (A.63) and is used as the weighting matrix in the criterion function. In the output

error method, the measurement noise is assumed known and the weighting function is arbitrary.

The first and second derivatives of the log-likelihood function with respect to unknown parameters are

$$\frac{\partial}{\partial \theta_j} \log(\mathscr{L}(\theta|z)) = - \sum_{i=1}^{N} \nu^T(i)\, \hat{R}^{-1} \frac{\partial \nu(i)}{\partial \theta_j} \tag{A.65}$$

$$\frac{\partial^2 \log(\mathscr{L}(\theta|z))}{\partial \theta_j\, \partial \theta_k} = - \sum_{i=1}^{N} \left\{ \frac{\partial \nu^T(i)}{\partial \theta_k} \hat{R}^{-1} \frac{\partial \nu(i)}{\partial \theta_j} + \nu^T(i)\, \hat{R}^{-1} \frac{\partial^2 \nu(i)}{\partial \theta_j\, \partial \theta_k} \right\} \tag{A.66}$$

The terms in the second derivative are approximated as

$$\frac{\partial^2 \log(\mathscr{L}(\theta|z))}{\partial \theta_j\, \partial \theta_k} = - \sum_{i=1}^{N} \left\{ \frac{\partial \nu^T(i)}{\partial \theta_k} \hat{R}^{-1} \frac{\partial \nu(i)}{\partial \theta_j} \right\} \tag{A.67}$$

No Measurement Noise

If all states are measured with no noise, the covariance of the error in state estimates is zero at the beginning of any time update,

$$\begin{aligned} &P(i-1|i-1) = 0 \\ \text{and} \quad &\hat{x}(i-1|i-1) = x(i-1) \end{aligned} \tag{A.68}$$

It is easy to show in this case that for fast sampling the log-likelihood function is quadratic in the difference between measured values of $\dot{x}$ and $f(x, u, \theta, t)$. The method reduces to the equation error method, the weight W being chosen as

$$W = \frac{1}{T} \int_{o}^{T} (\dot{x} - f(x,u,\theta,t))(\dot{x} - f(x,u,\theta,t))^T dt \qquad (A.69)$$

Thus, the maximum likelihood method and equation-error methods are equivalent except for the technique for choosing the weighting matrix.

APPENDIX B
EQUATIONS OF MOTION AND LIST OF PARAMETERS

Longitudinal Case

$$\dot{x} = Fx + Gu + \Gamma w$$

state vector
$$x = \begin{bmatrix} \alpha \\ u \\ q \\ \theta \\ \alpha_g \end{bmatrix} \begin{matrix} \text{(angle-of-attack perturbation)} \\ \text{(longitudinal velocity perturbation)} \\ \text{(pitch rate perturbation)} \\ \text{(pitch attitude perturbation)} \\ \text{(angle-of-attack due to gusts)} \end{matrix}$$

control vector
$$u = \begin{bmatrix} \delta_e \end{bmatrix} \quad \text{(horizontal elevator deflection from trim)}$$

$$F = \begin{bmatrix} Z_\alpha & Z_u & Z_q & -g/V \sin \theta_o & Z_\alpha \\ X_\alpha & X_u & X_q & -g \cos \theta_o & X_\alpha \\ M_\alpha & M_u & M_q & 0 & M_\alpha \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\omega_c \end{bmatrix}$$

$$G = \begin{bmatrix} Z_{\delta_e} \\ X_{\delta_e} \\ M_{\delta_e} \\ 0 \\ 0 \end{bmatrix} \qquad \Gamma w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ n_6 \end{bmatrix}$$

where

g is gravitational acceleration,

V is steady-state, total velocity,

θ_o is steady-state pitch angle

measurements $y = Hx + Du + v$

$$
Hx+Du = \begin{bmatrix} K_\alpha(\alpha+\alpha_q) + \frac{\ell_\alpha K_\alpha}{V} q \\ u \\ q \\ \theta \\ -V\left[Z_\alpha(\alpha+\alpha_q)+Z_{\delta_e}'\delta_e\right] \end{bmatrix} \qquad v = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{bmatrix}
$$

where

K_α = angle-of-attack vane scale factor

ℓ_α = longitudinal distance of angle-of-attack vane forward of c.g.

Lateral Case

$$
\text{state vector} \quad x = \begin{bmatrix} p \\ r \\ \beta \\ \phi \\ \beta_g \end{bmatrix} \begin{array}{l} \text{(roll rate)} \\ \text{(yaw rate)} \\ \text{(sideslip angle)} \\ \text{(roll angle)} \\ \text{(sideslip angle due to gusts)} \end{array}
$$

$$\text{control vector} \quad u = \begin{bmatrix} \delta_a \\ \delta_r \end{bmatrix} \quad \begin{matrix} \text{(aileron deflection)} \\ \text{(rudder deflection)} \end{matrix}$$

$$F = \begin{bmatrix} L_p & L_r & L_\beta & 0 & L_\beta \\ N_p & N_r & N_\beta & 0 & N_\beta \\ \sin\alpha_o & -\cos\alpha_o & Y_\beta & g/V\cos\theta_o & Y_\beta \\ 1 & \tan\theta_o & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\omega_c \end{bmatrix}$$

$$G = \begin{bmatrix} L_{\delta_a} & L_{\delta_r} \\ N_{\delta_a} & N_{\delta_r} \\ Y_{\delta_a} & Y_{\delta_r} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \Gamma w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ n_6 \end{bmatrix}$$

where θ_o is steady-state pitch angle.

measurements $y = Hx + Du + v$

$$Hx+Du = \begin{bmatrix} p \\ r \\ K_\beta(\beta+\beta_g) + \frac{\ell_\beta K_\beta}{V} r \\ \phi \\ V\left[Y_\beta(\beta+\beta_g)+Y_{\delta_a}\delta_a+Y_{\delta_r}\delta_r\right] \end{bmatrix} \qquad v = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{bmatrix}$$

where K_β and ℓ_β are the sideslip vane's scale factor and longitudinal distance forward of the c.g., respectively.

LONGITUDINAL PARAMETERS

NUMBER	PARAMETER	UNITS
1	Z_α	sec^{-1}
2	Z_u	ft^{-1}
3	Z_q	----
4	γ_o	rad
5	X_α	$ft\ sec^{-2}$
6	X_u	sec^{-1}
7	X_q	$ft\ sec^{-1}$
8	M_α	sec^{-2}
9	M_u	$ft^{-1}\ sec^{-1}$
10	M_q	sec^{-1}
11	ω_c	$rad\ sec^{-1}$
12	Z_{δ_e}	sec^{-1}
13	X_{δ_e}	$ft\ sec^{-2}$
14	M_{δ_e}	sec^{-2}
15	Z_o	sec^{-1}
16	X_o	$ft\ sec^{-2}$
17	M_o	sec^{-2}
18	θ_o	sec^{-1}
19	K_α	----
20	$\ell_\alpha K_\alpha/V$	sec
21	Q	$rad^2\ sec^{-1}$
22	R_α	rad^2
23	R_u	$ft^2\ sec^{-2}$
24	R_q	$rad^2\ sec^{-2}$
25	R_θ	rad
26	R_{a_z}	$ft^2\ sec^{-4}$

LATERAL PARAMETERS

NUMBER	PARAMETER	UNITS
1	L_p	sec^{-1}
2	L_r	sec^{-1}
3	L_β	sec^{-2}
4	N_p	sec^{-1}
5	N_r	sec^{-1}
6	N_β	sec^{-2}
7	α	rad
8	Y_β	ft sec^{-2}
9	θ	rad
10	ω_c	rad sec^{-1}
11	L_{δ_a}	sec^{-2}
12	N_{δ_a}	sec^{-2}
13	Y_{δ_a}	sec^{-1}
14	L_{δ_r}	sec^{-2}
15	N_{δ_r}	sec^{-2}
16	Y_{δ_r}	sec^{-1}
17	L_o	Unassigned - fill in any value
18	N_o	Unassigned - fill in any value
19	Y_o	Unassigned - fill in any value
20	ϕ_o	Unassigned - fill in any value
21	Q	rad^2 sec^{-1}
22	K_β	
23	$\ell_\beta K_\beta/V$	sec
24	R_p	sec^{-2}
25	R_r	sec^{-2}
26	R_β	rad^2
27	R_ϕ	rad^2
28	R_{a_y}	ft^2 sec^{-2}

REFERENCES

[1] Hall, W.E., Gupta, N.K., and Smith, R.G., "Identification of Aircraft Stability and Control Derivatives for the High Angle-of-Attack Regime," Final Report to ONR on Contract N00014-72-C-0328, March 1974.

[2] Gupta, N.K., and Mehra, R.K., "Computational Methods in Maximum Likelihood Estimation and Reduction in Sensitivity Function Computations," to appear in IEEE Trans. Automatic Control, December 1974.

[3] Gupta, N.K., and Hall, W.E., "Input Design for Identification of Aircraft Stability and Control Derivatives," to appear NASA CR, September 1974.

[4] Segall, I., Mohr, R., and Gupta, N.K., "User's Manual for SCIDNT I," prepared for NATC on Contract N00014-72-C-0328, September 1974.

* Not identified

[All in units of ft., sec., rad.)

[Data Supplied by NATC to SCI

* Not identified

[All in units of ft., rad., sec.)

[Data Supplied by NATC TO SCI - July, 1974]

* Not identified

[All in units of ft., rad., sec.)

[Data Supplied by NATC TO SCI - April, 1974]

Table 3.3

F-14 RUN2

Altitude 30,000 ft.
Speed 362 knots
Sweepback 20°

DERIVATIVE	STARTING VALUE (LEAST SQUARES AND WIND TUNNEL)	ESTIMATED RAW DATA	VALUE (WITH 1σ BOUND) OUTLIERS REMOVED
Z_α	-0.38	-0.347 (.033)	-0.339 (.00271)
Z_u	-0.00066	*	*
Z_q	1.04	*	*
γ_o	0.0	*	*
X_α	285.0	*	*
X_u	-3.308	*	*
X_q	42.34	*	*
M_α	-0.755	-1.276 (.0789)	-1.285 (.00649)
M_u	0.00416	*	*
M_q	-0.600	-0.736 (.139)	-0.7068 (.011)
ω_c	0.50	*	*
$Z\delta_e$	-0.20	0.0378 (.218)	0.119 (.020)
$X\delta_e$	0.0	*	*
$M\delta_e$	-5.46	-5.357 (.452)	-5.338 (.033)
Q	0.0	*	*
k_α	1.0	*	*
$\frac{k\alpha\ell\alpha}{V}$	0.0	*	-0.0754 (.0062)
b_α	0.0	*	0.000124
b_u	0.0	*	-0.0934
b_q	0.0	*	0.00056
b_θ	0.0	*	0.000213
ba_z	0.0	*	-0.05
σ_α	0.0075	0.0428	0.00159
σ_u	100.0	52.0	4.72
σ_q	0.007	0.0424	0.0038
σ_θ	0.015	0.0877	0.0021
σ_{a_z}	0.9	8.1	0.77

Table 3.4

RUN1 (DOUBLET INPUT)

Altitude 30,000 ft.
Speed 360 knots
Sweepback 20°

PARAMETER	PARAMETER VALUE (WITH 1σ ERROR BOUND IN BRACKETS)			
	STARTING VALUE (FROM PREVIOUS RUN)	q FROM DATA (INCORRECT)	q NOT USED	q CORRECTED AND USED
Z_α	-0.3394	0.545	-0.4268 (.0097)	-0.4089 (.0059)
Z_u	-0.00066	*	*	*
Z_q	1.04	*	*	*
γ_o	0.0	*	*	*
X_α	285.0	*	*	*
X_u	-3.3	*	*	*
X_q	42.3	*	*	*
M_α	-1.28	-4.59	-2.002 (.018)	-1.88 (.0078)
M_u	0.0042	*	*	*
M_q	-0.74	-2.73	-0.4907 (.024)	-0.6726 (.0089)
ω_c	0.5	*	*	*
Z_{δ_e}	0.119	0.639	0.0856 (.014)	-0.0147 (.0136)
X_{δ_e}	0.0	*		
M_{δ_e}	5.34	0.416	4.154 (.093)	5.30 (.033)
Q	0.0	*	*	*
k_α	1.0	*	*	0.881 (.0089)
$\frac{k_\alpha \ell_\alpha}{V}$	0.0	-0.08	-0.0594 (.0066)	*
b_α	0.0	-0.0328	0.00609	-0.00071
b_u	0.0	-2.06	0.344	-0.147
b_q	0.0	-0.0104	0.00164	-0.000877
b_θ	0.0	-0.0378	0.0112	-0.00505
b_{a_z}	0.0	-0.81	0.922	-0.528
σ_α	0.0075	0.034	0.0072	0.00262
σ_u	100.0	2.7	4.05	1.35
σ_q	0.007	0.023	0.054	0.0034
σ_θ	0.015	0.063	0.0131	0.0100
σ_{a_z}	0.9	10.0	2.6	2.1

Table 3.5

LATERAL MOTIONS (RUN8)*

Altitude 30,000 ft.
Speed 360 knots
Sweepback 20°

PARAMETER	STARTING VALUE (LEAST SQUARES)	MAXIMUM LIKELIHOOD ESTIMATE (WITH 1σ BOUND)	
		NO PROCESS NOISE	PROCESS NOISE
L_p	-1.975	-2.004 (.0235)	-1.948 (.0205)
L_r	4.877	2.686 (.0645)	2.369 (.0541)
L_β	-8.29	-8.500 (.0982)	-6.952 (.0846)
N_p	-0.004	*	*
N_r	-0.62	*	*
N_β	1.96	*	*
Y_β	-0.05	-0.0649 (.00190)	-0.0819 (.00121)
α_o	0.0	*	*
θ_o	0.0	*	*
L_{δ_a}	10.29	*	*
N_{δ_a}	0.24	*	*
Y_{δ_a}	0.0	*	*
L_{δ_r}	-2.828	-3.337 (.0865)	-1.446 (.0661)
N_{δ_r}	1.55	*	*
Y_{δ_r}	0.42	-0.015 (.0018)	-0.0150 (.00121)
K_β	1.0	*	*
$K_\beta \ell_\beta / V$	0.03	*	*
ω_c	0.5	*	*
Q	0.0	*	0.000347 (.000026)
b_p	-	-0.000209	-0.00264
b_r	-	-0.000126	0.000318
b_β	-	-0.0000590	0.000636
b_ϕ	-	-0.00230	-0.00516
b_{a_y}	-	0.00568	-0.0629
σ_p	0.005	0.0405	0.0246
σ_γ	0.005	0.0196	0.0069
σ_β	0.005	0.0190	0.00487
σ_ϕ	0.005	0.0192	0.0245
σ_{a_y}	0.5	1.57	0.85